

STIC-Biotech/ChemLib

From: Helmer, Georgia  
Sent: Tuesday, October 08, 2002 6:17 PM  
To: STIC-Biotech/ChemLib  
Subject: Seq Search - 09/289,346

77501

Q 42-55

In case - 09/289,346:

could you please do a sequence search and an oligo search on SEQ ID Nos 1 thru 10.

Both inhouse and commercial databases. And also an Interference search.

Thank you in advance for your help.

Georgia L. Helmer Ph.D.  
Patent Examiner  
Crystal Mall 1, 9D14  
AU 1638  
703-308-7023  
Georgia.Helmer@USPTO.gov  
mailbox 9e12

Seq ID #1 = wt

RECEIVED  
OCT - 9 2002  
STIC

Point of Contact  
P. Sheppard  
Telephone number: (703) 308-4499

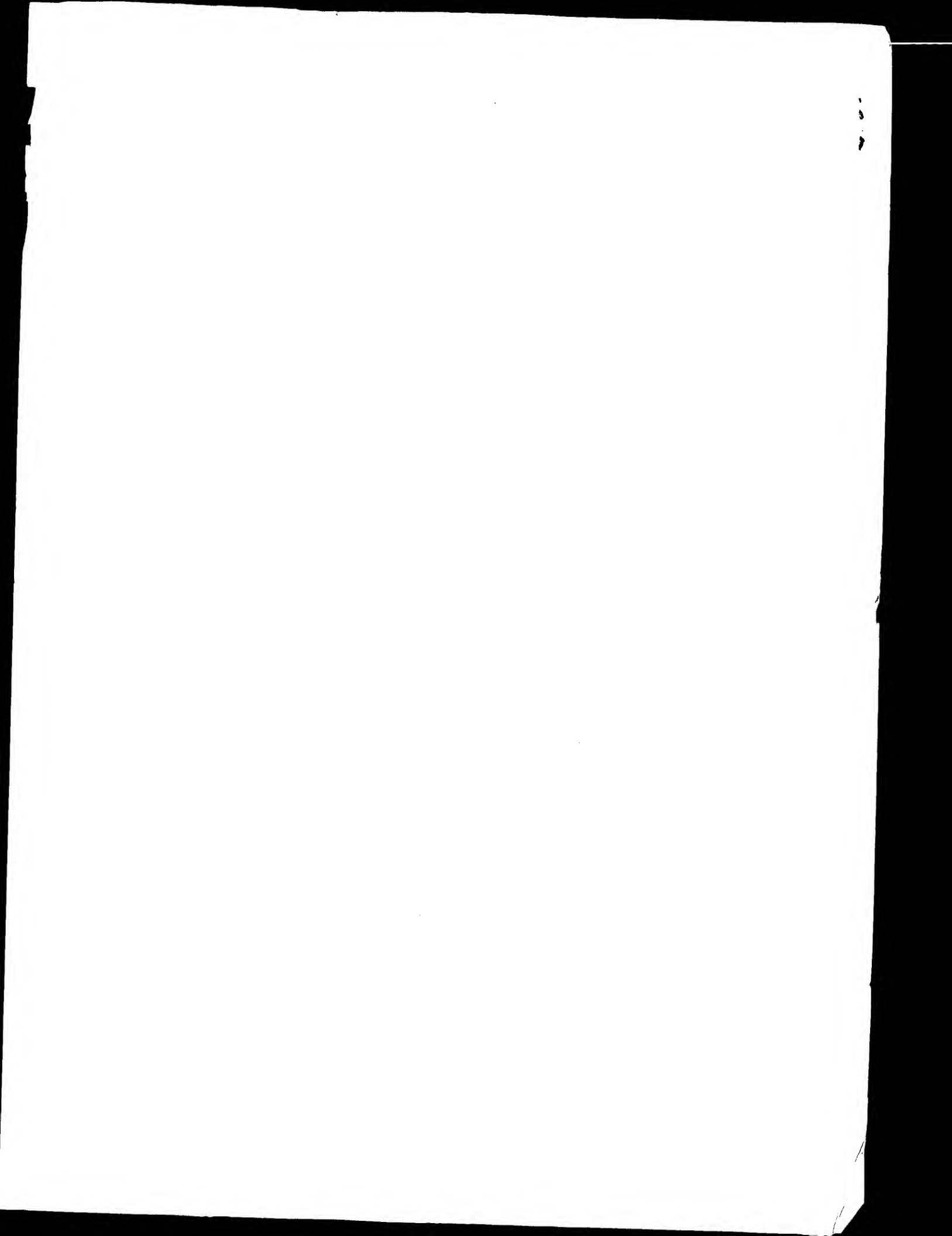
Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 10/11/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_





GenCore version 4.5  
1993 - 2000 Compugen Ltd

Copyright (c) 1999

OM protein - protein search, using 5w model  
 Run on: october 10, 2002, 01:43:47 ; Search time 118.38 Seconds  
 (without alignments)  
 102.295 Million cell updates/sec

Title:	US-09-289-346A-1
Perfect score:	364
Sequence:	1 FLWGEFQVDCRSRGCGQF.....PQFHLNLSNLDRIFDKTPR 70
Scoring table:	BLOSUM62
	Gapop 10.0, Gapext 0.5
	30,000 residues

Searched: 304222  
 Choosing parameters:

```

Total number of records: 1000000
Minimum DB seq length: 2000000000
Maximum DB seq length: 2000000000
post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database :

```

1:  sp:archaea:*
2:  sp:bacteria:*
3:  sp:fungi:*
4:  sp:human:*
5:  sp:invertebrate:*
6:  sp:mammal:*
7:  sp:mhc:*
8:  sp:organelle:*
9:  sp:phase:*
10: sp:plant:*
11: sp:podent:*
12: sp:yirus:*
13: sp:vertebrate:*
14: sp:unclassified:*
15: sp:yirus:*
16: sp:bocteria:*
17: sp:archaeo:*

```

```

17: sp_cheap:
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.
SUMMARIES

```

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	315	86.5	351	12	Q9B10	09i10 tomato seve
2	313	86.0	352	12	Q9E000	09e000 tomato rugo
3	296	81.5	226	12	Q9W727	09w727 leonurus mo
4	296	81.5	226	12	Q9WHF6	09whf6 tomato mild
5	294	80.8	361	12	Q67574	067574 bean golden
6	292	80.2	325	12	Q90DB1	09e1b1 cowpea gold
7	292	80.2	344	12	Q9ELT8	09elt8 sweet potat
8	292	80.2	364	12	Q9C555	09c555 sweet potat
9	289	79.4	149	12	Q9E693	09e693 sida golden
10	281	77.2	233	12	Q9YL44	09yl44 macroptilium
11	275	75.5	234	12	Q93180	09y1a4 geminivirid
12	275	74.7	190	12	Q9Z084	03180 tobacco lea
13	272	74.5	190	12	Q9Z084	09z084 tobacco lea
14	271	74.5	190	12	Q9W827	09w827 tobacco lea
15	271	74.5	208	12	Q9Z0C4	09z0c4 tobacco lea

17	271	74.5	359	12	Q91M88	Q92C00 tobacco lea
18	269	73.9	208	12	Q920C0	Q920B8 tobacco lea
19	269	73.9	208	12	Q920B8	Q920B3 tobacco lea
20	267	73.4	203	12	Q920B6	Q73577 cotton leaf
21	266	73.1	203	12	Q920B3	Q91B66 ageratum ye
22	266	73.1	363	12	Q73577	Q96X10 ageratum ye
23	261	71.7	359	12	Q91B86	Q920A0 tobacco lea
24	261	71.7	308	12	Q92DX10	Q920A5 cotton leaf
25	259	71.2	208	12	Q72705	Q72719 cotton leaf
26	259	71.2	363	12	Q72719	Q920B6 tobacco lea
27	258	70.9	190	12	Q920C6	Q920C6 tobacco pseu
28	258	70.9	208	12	Q88888	Q88888 tomato mos
29	258	70.6	349	12	Q56816	Q56816 chayote mos
30	257	70.6	362	12	Q92Y24	Q92Y24 tomato yell
31	257	70.6	359	12	Q92Y24	Q92Y22 tomato yell
32	256	70.3	359	12	Q92Y22	Q92Y27 tomato yell
33	256	70.3	359	12	Q9Y0X7	Q91127 macroptilium
34	256	70.3	359	12	Q9Y1Z7	Q72892 beet curly
35	256	70.3	348	12	Q91M55	Q65438 beet curly
36	255	70.1	353	12	Q72692	Q88942 tomato yell
37	255	70.1	353	12	Q65438	Q88942 tomato lea
38	255	70.1	359	12	Q88942	Q920A7 tobacco lea
39	254	69.8	190	12	Q920A7	Q91A48 okra enatio
40	254	69.8	362	12	Q91A48	Q91E17 cotton leaf
41	253	69.5	307	12	Q91E77	Q92723 cotton leaf
42	253	69.5	361	12	Q72723	Q96620 african tom
43	252	69.2	231	12	Q96620	Q91n24 south africa
44	252	69.2	354	12	Q91N24	

## ALIGNMENTS

[illegible]

AC 09E000;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DE 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 GN REPLICATION-ASSOCIATED PROTEIN.  
 OS AC1.  
 OC Tomato rugose mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN NCBI\_TaxID=134599;  
 RP SEQUENCE FROM N.A.  
 RA Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,  
 RT "Molecular Cloning and Characterization of Tomato Rugose mosaic virus  
 (TRMV), a Begomovirus isolated from Tomatoes at Itangulo Mineiro,  
 Minas Gerais, Brazil.",  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF291705; AAG15546.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 SQ SEQUENCE 352 AA: 40012 MW: 47CD5838E24D613 CRC64;

## Query Match

Best Local Similarity 86.0%; Score 313; DB 12; Length 352;  
 Matches 58; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 1 TIVGCEFOVDRSARGCOTSNDAAEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60  
 Db 111 TIVGCEFOVDRSARGCOTSNDAAEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60  
 OY 61 DRIEDKTPPEP 70  
 Db 171 DRIEFAKPEP 180

RESULT 3  
 ID 009727 PRELIMINARY; PRT: 226 AA.  
 AC 009727;  
 DT 01-JUL-1997 (TREMblrel. 04, Created)  
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
 DE REP. PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonturus mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN NCBI\_TaxID=58177;  
 RP SEQUENCE FROM N.A.  
 RA STRAIN:LEWY-BRAZIL 1;  
 RT "Variability in geminiviruses associated with Phaseolus vulgaris in  
 Brazil.",  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U92532; AAB51157.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 FT NON\_TER  
 SQ SEQUENCE 226 AA: 25617 MW: 73CDB6E76083FC5 CRC64;

## Query Match

Best Local Similarity 81.3%; Score 296; DB 12; Length 226;  
 Matches 57; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 TIVGCEFOVDRSARGCOTSNDAAEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60  
 Db 111 TIVGCEFOVDRSARGCOTSNDAAEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60

OY 61 DRIEDKTPPEP 70  
 Db 171 DRIEFAKPEP 180

RESULT 4  
 ID 09WHF6 PRELIMINARY; PRT: 226 AA.  
 AC 09WHF6;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Tomato mild mottle geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
 RN NCBI\_TaxID=92943;  
 RP SEQUENCE FROM N.A.  
 RA STRAIN:HN96-H5KW;  
 RA Nakha M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
 RT "Molecular characterization and DNA-based detection methods for  
 vegetable-infecting geminiviruses in Central America.",  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF131071; AAD33471.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 FT NON\_TER  
 SQ SEQUENCE 226 AA: 25941 MW: 2EA4116712871A23 CRC64;

## Query Match

Best Local Similarity 81.3%; Score 296; DB 12; Length 226;  
 Matches 54; Conservative 12; Mismatches 4; Indels 0; Gaps 0;

OY 1 TIVGCEFOVDRSARGCOTSNDAAEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60  
 Db 111 TIVGCEFOVDRSARGCOTSNDAAEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60  
 OY 61 DRIEDKTPPEP 70  
 Db 171 DRIEFAKPEP 180

RESULT 5  
 ID 067574 PRELIMINARY; PRT: 361 AA.  
 AC 067574;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN BEAN.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN NCBI\_TaxID=10839;  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 RT "Cloning of the complete DNA genomes of four bean-infecting  
 geminiviruses and determining their infectivity by electric discharge  
 particle acceleration.",  
 RL Phytopathology 81:980-985(1991).  
 SQ SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 RT "Differentiation of bean-infecting geminiviruses by nucleic acid

Thu Oct 10 16:05:31 2002

us-09-289-346a-1.rspt

Page 3

```

RT hybridization probes and aspects of bean golden mosaic in Brazil."
RL Plant Dis. 75:336-342(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA Gilbertson R.L., Farla J.C., Ahlquist P.G., Maxwell D.P.;
RT "Genetic diversity in geminiviruses causing bean golden mosaic
RT disease: The nucleotide sequence of the infectious cloned DNA
RT components of a Brazilian isolate of bean golden mosaic geminivirus."
RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
RL EMBL: M86866; AAA46312.1;
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR Prodom: PD000736; Gemini_AL1.
SO SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 80.8%; Score 294; DB 12; Length 361;
Best Local Similarity 80.6%; Pred. No. 6.8e-25;
Matches 54; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

OY 4 WGEFOVDGRSARGCQTSNDAAAEALNASKKEALQIIREKIPKYLFOFHNLNSLDR1 63
DB 113 WGHFOVDGRSARGCQOTINDAASEALNASKKEALQIIREKIPKYLFOFHNLNSLDR1 172
OY 113 WGHFOVDGRSARGCQOTINDAASEALNASKKEALQIIREKIPKYLFOFHNLNSLDR1 172
DB 173 FTKKPP 179

RESULT 6
ID 09QDB1 PRELIMINARY; PRT; 225 AA.
AC 09QDB1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN RRP.
OS cowpea golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=69263;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CGM-VR;
RC Farla J.C.;
RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT Brazil."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF188708; AF06318.1;
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR Prodom: PD000736; Gemini_AL1.
FT NON_TER 225
SO SEQUENCE 225 AA; 25766 MW; 1089CB6BD8D15B5D CRC64;

Query Match 80.2%; Score 292; DB 12; Length 225;
Best Local Similarity 80.6%; Pred. No. 6.8e-25;
Matches 54; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

OY 4 WGEFOVDGRSARGCQTSNDAAAEALNASKKEALQIIREKIPKYLFOFHNLNSLDR1 63
DB 113 WGHFOVDGRSARGCQOTINDAASEALNASKKEALQIIREKIPKYLFOFHNLNSLDR1 172
OY 113 WGHFOVDGRSARGCQOTINDAASEALNASKKEALQIIREKIPKYLFOFHNLNSLDR1 172
DB 173 FTKKPP 179

RESULT 7
ID 09ELT8
```

```

ID 09ELT8 PRELIMINARY; PRT; 314 AA.
AC 09ELT8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 19, Last annotation update)
DE REPLICATION ASSOCIATION PROTEIN.
GN AC1.
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=100755;
RN [1]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;
RT "Detection of a geminivirus infecting sweet potato in the United
RT States."
RT Plant Dis. 82:1253-1257(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;
RT "Detection of a geminivirus infecting sweet potato in the United
RT States."
RL Plant Dis. 82:1253-1257(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A.;
RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT virus: nucleotide sequence and phylogenetic relationships."
RL Online Publication.
DR EMBL: AF104036; AAD47173.1;
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR Prodom: PD000736; Gemini_AL1.
SO SEQUENCE 364 AA; 40680 MW; 5E79752431A09D6E CRC64;

Query Match 80.2%; Score 292; DB 12; Length 314;
Best Local Similarity 83.8%; Pred. No. 9.7e-25;
Matches 57; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 TIVWGEFOVDGRSARGCQTSNDAAAEALNASKKEALQIIREKIPKYLFOFHNLNSL 60
DB 110 TIVWGEFOVDGRSARGCQOTINDAAEALNASKKEALQIIREKIPKYLFOFHNLNSL 169
OY 110 TIVWGEFOVDGRSARGCQOTINDAAEALNASKKEALQIIREKIPKYLFOFHNLNSL 169
DB 170 DRIFSP 177

RESULT 8
ID 09Q555 PRELIMINARY; PRT; 364 AA.
AC 09Q555;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 19, Last annotation update)
DE REPLICATION INITIATION PROTEIN AC1.
GN AC1.
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=100755;
RN [1]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;
RT "Detection of a geminivirus infecting sweet potato in the United
RT States."
RL Plant Dis. 82:1253-1257(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A.;
RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT virus: nucleotide sequence and phylogenetic relationships."
RL Online Publication.
DR EMBL: AF104036; AAD47173.1;
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR Prodom: PD000736; Gemini_AL1.
SO SEQUENCE 364 AA; 40680 MW; 5E79752431A09D6E CRC64;
```

Query Match  
Best Local Similarity 80.2%; Score 292; DB 12; Length 364;  
Matches 57; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 60  
DB 110 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 169  
OY 61 DRIFPKTPEP 68  
DB 170 DRIFSKPEP 177

RESULT 9  
ID 098693 PRELIMINARY; PRT: 185 AA.  
AC 098693;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
GN AC1.  
OS sida golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51034;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA;  
RA Roye M.E.; McLaughlin W.A.; Nakha N.K.; Maxwell D.P.;  
RT "Genetic Diversity among geminiviruses associated with the weed  
RT species sida spp, Macroptilium lathyroides, and Wissadula amplissima  
RL plant Dis. 81:1251-1258(1997).  
DR EMBL: U67926; AAB97865.1;  
DR InterPro: IPR001191; Geminl\_AL1.  
DR Pfam: PF00799; Geminl\_AL1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR ProDom: PD000736; Geminl\_AL1; 1.  
FT NON\_TER 1 1  
ET NON\_TER 185 185  
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match  
Best Local Similarity 79.4%; Score 289; DB 12; Length 185;  
Matches 52; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

OY 1 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 60  
DB 89 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 148  
OY 61 DRIFPKTPEP 70  
DB 149 DRIFSKPEP 158

RESULT 10  
ID P88975 PRELIMINARY; PRT: 149 AA.  
AC P88975;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
GN AC1.  
OS Macroptilium golden mosaic geminivirus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA;  
RA Roye M.E.;  
RL Thesis (1996), Biochemistry, University of the West Indies, Jamaica.

DR EMBL: U75278; AAB36919.1;  
DR InterPro: IPR001191; Geminl\_AL1.  
DR Pfam: PF00799; Geminl\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR ProDom: PD000736; Geminl\_AL1; 1.  
FT NON\_TER 1 1  
ET NON\_TER 149 149  
SQ SEQUENCE 149 AA; 16785 MW; E4CF5ED4C9C0508 CRC64;

Query Match  
Best Local Similarity 77.2%; Score 281; DB 12; Length 149;  
Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

OY 1 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 60  
DB 52 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 111  
OY 61 DRIFPKTPEP 70  
DB 112 DRIFMKDPEP 121

RESULT 11  
ID 09Y1A4 PRELIMINARY; PRT: 233 AA.  
AC 09Y1A4;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
GN REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
OS Macroptilium golden mosaic geminivirus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA STRAIN 1;  
RA Roye M.E.;  
RT "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses  
RT from Jamaica."  
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA STRAIN 1;  
RA Roye M.E.; McLaughlin W.A.; Maxwell D.P.;  
RT "Molecular characterization of two distinct geminiviruses infecting M.  
RT lathyroides from Jamaica."  
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF098940; AAD17850.1;  
DR InterPro: IPR001191; Geminl\_AL1.  
DR Pfam: PF00799; Geminl\_AL1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR ProDom: PD000736; Geminl\_AL1; 1.  
FT NON\_TER 233 233  
SQ SEQUENCE 233 AA; 26356 MW; AA490AF4D2166A02 CRC64;

Query Match  
Best Local Similarity 77.2%; Score 281; DB 12; Length 233;  
Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

OY 1 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 60  
DB 110 TLWGEFQVYDGRSARGGQTSNDAAAEALNASSKEEALQITREKPEKYLFOFHNLSNL 169  
OY 61 DRIFPKTPEP 70  
DB 170 DRIFMKDPEP 179

RESULT 12  
039180

ID 039180 PRELIMINARY; PRT: 234 AA.  
 AC 039180;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE AC1 (FRAGMENT).  
 OS Geminiviridae.  
 OC Viruses; ssDNA viruses.  
 NCBI\_TaxID=10811;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TOMATO STRAIN:  
 RA Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;  
 RT "Partial Characterization of Two Whitefly-Transmitted Geminiviruses  
 Infecting Tomatoes in Venezuela."  
 RL Plant Dis. 81:312-312(1997).  
 DR EMBL: AF026553; AAB82605.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR PRODOM: PD000736; GEMCOATAL1.  
 DR NON\_TER 234 234  
 FT SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;  
 SQ

Query Match 75.5%; Score 275; DB 12; Length 234;  
 Best Local Similarity 71.4%; Pred. No. 5.5e-23;  
 Matches 50; Conservative 13; Mismatches 7; Indels 0; Gaps 0;

OY 1 TLVWGEFOVDRSARGCCQTSNDAAEALNASSKEALQITREKIPKFLPFQHNLSNL 60  
 DB 110 TLEMGTFQIDGRSARGCGQOTANDAAEALNASSKEALQITREKIPKFLPFQHNLSNL 169  
 OY 61 DRIFDKTPEP 70  
 DB 170 DRIFDKAPNP 179

RESULT 13  
 O92089 PRELIMINARY; PRT: 190 AA.  
 AC 092089;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL  
 AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 1 1  
 FT SEQUENCE 190 AA; 21432 MW; AAC093D1D1610FAD CRC64;  
 SQ

Query Match 74.7%; Score 272; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 9.5e-23;  
 Matches 53; Conservative 8; Mismatches 8; Indels 16; Gaps 1;  
 OY 1 TLVWGEFOVDRSARGCCQTSNDAAEALNASSKEALQITREKIPKFLPFQHNLSNL 60  
 DB 110 TLEMGTFQIDGRSARGCGQOTANDAAEALNASSKEALQITREKIPKFLPFQHNLSNL 169

DB 85 TLEMGTFQIDGRSARGCCQTSNDAAEALNASSKEALQITREKIPKFLPFQHNLSNL 144  
 OY 61 DRIFDKTPEP 70  
 DB 145 DRIFAPLEVEVCPFTASSFDQVPE 169

RESULT 14  
 O92084 PRELIMINARY; PRT: 190 AA.  
 AC 092084;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-2, PARTIAL  
 AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001318; BAA34039.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 1 1  
 FT SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;  
 SQ

Query Match 74.7%; Score 272; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 9.5e-23;  
 Matches 53; Conservative 8; Mismatches 8; Indels 16; Gaps 1;

OY 1 TLVWGEFOVDRSARGCCQTSNDAAEALNASSKEALQITREKIPKFLPFQHNLSNL 60  
 DB 85 TLEMGTFQIDGRSARGCGQOTANDAAEALNASSKEALQITREKIPKFLPFQHNLSNL 144  
 OY 61 DRIFDKTPEP 70  
 DB 145 DRIFAPLEVEVCPFTASSFDQVPE 169

RESULT 15  
 O9W827 PRELIMINARY; PRT: 190 AA.  
 AC 09W827;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE CORA1-1, PARTIAL AND  
 COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CORA1;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001303; BAA34010.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.

Thu Oct 10 16:05:31 2002

us-09-289-346a-1.rsp

Page 6.

FT	NON_TER	1	1
FT	NON_TER	190	190
SO	SEQUENCE	190 AA;	21444 MW; 93C3742A8EBDB7EB CRC64;

Query Match	74.58;	Score 271;	DB 12;	Length 130;
Best Local Similarity	62.48;	Pred. No. 1.2e-22;		
Matches 53;	Conservative 8;	Mismatches 8;	Indels 16;	Gaps 1;

```

QY      1  TIWGEFQVDDGFSRAGCCGOTSNDAAALALNLSKSEKALQITREKPEIKYLFQFHLNSNL 60
Db      85  TLEMGTFQVDDGFSRAGCCGONANADCAELALNLSKSEKALSTIREKLPKDFIYQHLNSNL 144
QY      61  DRI-----FKTPE 69
Db     145  DRIFAPLEFVCPFSNSSFQVQPE 169

```

Search completed: October 10, 2002, 02:01:59  
Job time: 1092 sec



```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 protein
OS Potato yellow mosaic virus (isolate Venezuela)
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311403; PubMed=1866690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RT potato yellow mosaic virus."
RL J. Gen. Virol. 72:1515-1520(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00940; BAA00782.1;
DR PIR: J00364; OOCVPT.
DR InterPro: IPR001191; Gemin1_AL1.
DR Pfam: PF00799; Gemin1_AL1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemin1_AL1; 1.
KW ATP-binding.
FT NP-BIND.
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match
Best Local Similarity 72.58; Score 264; DB 1; Length 361;
Matches 49; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60
DB 110 TLWGEFQVDRSARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 169
QY 61 DRIDKPEP 69
DB 170 DRIDKPEP 178

RESULT 3
VAL1_CLV
ID VAL1_CLV STANDARD; PRT; 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Gay M.R.;
RT "Nucleotide sequence of cassava latent virus DNA."
RL Nature 301:260-262(1983).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02057; NOT_ANNOTATED_CPS.
DR InterPro: IPR001191; Gemin1_AL1.
DR Pfam: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemin1_AL1; 1.
KW ATP-binding.
FT NP-BIND.
SQ SEQUENCE 358 AA; 40346 MW; ED173E753EB92069 CRC64;

Query Match
Best Local Similarity 69.08; Score 251; DB 1; Length 358;
Matches 45; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60
DB 109 TLWGEFQVDRSARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 168
QY 61 DRIDKPEP 70
DB 169 DRIDKPEP 178

RESULT 4
VAL1_CLV
ID VAL1_CLV STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain)."
RL Nucleic Acids Res. 18:197-198(1990).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17095; CA34953.1;
DR PIR: S07594; S07594.
DR InterPro: IPR001191; Gemin1_AL1.
DR Pfam: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemin1_AL1; 1.
KW ATP-binding.
FT NP-BIND.
SQ SEQUENCE 358 AA; 40435 MW; 1DB16B0CB2D5E2C CRC64;

Query Match
Best Local Similarity 69.08; Score 251; DB 1; Length 358;
Matches 45; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60
DB 109 TLWGEFQVDRSARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 168

```



OY 61 DRIFDKTPEP 70  
 DB 169 DRIFDQPPAP 178

RESULT 5  
 VAL1\_TYLCV STANDARD: PRT: 362 AA.  
 AC 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN Tomato yellow leaf curl virus (strain Australia) (TYLCV).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID=36447;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=93139778; PubMed=8423446;  
 RA Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;  
 RT "Nucleotide sequence and genome organization of tomato leaf curl  
 geminivirus.";  
 RT J. Gen. Virol. 74:147-151(1993).  
 RL -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC PIR: J01887; J01887; Gemin1\_AL1.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 DR ATP-binding.  
 KW NP\_BIND 221 228 ATP (POTENTIAL).  
 FT SEQUENCE 362 AA: 41197 MW: 34357184B4704098 CRC64;  
 SQ

Query Match 67.0%; Score 244; DB 1; Length 362;  
 Best local Similarity 56.5%; Pred. No. 5; 1e-20;  
 Matches 48; Conservative 9; Mismatches 12; Indels 16; Gaps 1;

OY 1 TLVWGEFQVDGSRAGCGCOTSDNAAAEALNNSKEEALQIREKIPKYLFOFHINSLN 60  
 DB 110 TLVWGEFQVDGSRAGCGCOTSDNAAAEALNNSKEEALQIREKIPKYLFOFHINSLN 169  
 OY 61 DRI-----FDKTPPE 69  
 DB 170 DRIFDQPPAP 178

RESULT 6  
 VAL1\_TYLCV STANDARD: PRT: 359 AA.  
 AC 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID=37139;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=94256836; PubMed=8198442;  
 RA Morris E., Hidalgo E., Accotto G., Moriones E.;  
 RT "High similarity among the tomato yellow leaf curl virus isolates  
 from the west Mediterranean basin: the nucleotide sequence of an  
 infectious clone from Spain.";  
 RT Arch. Virol. 135:165-170(1994).  
 RL -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z25751; CAA81026.1; -  
 DR PIR: S39211; S39211.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 DR ATP-binding.  
 KW NP\_BIND 221 228 ATP (POTENTIAL).  
 FT SEQUENCE 359 AA: 41065 MW: 2D170A51EF80A3EC CRC64;  
 SQ

Query Match 65.1%; Score 237; DB 1; Length 359;  
 Best local Similarity 62.3%; Pred. No. 3; 1e-19;  
 Matches 43; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

OY 2 TLVWGEFQVDGSRAGCGCOTSDNAAAEALNNSKEEALQIREKIPKYLFOFHINSLN 61  
 DB 111 TLVWGEFQVDGSRAGCGCOTSDNAAAEALNNSKEEALQIREKIPKYLFOFHINSLN 170  
 OY 62 DRIFDKTPEP 70  
 DB 171 RVFOVPPAP 179

RESULT 7  
 VAL1\_PHVU STANDARD: PRT: 349 AA.  
 AC 006923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN Pepper huasteco virus (PHV).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID=28349;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=94015007; PubMed=8409944;  
 RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,  
 RT "Complete nucleotide sequence of pepper huasteco virus: analysis and  
 RT comparison with bipartite geminiviruses.";  
 RT J. Gen. Virol. 74:2225-2231(1993).  
 RL -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X70418; CAA9856.1; -  
 DR PIR: S31875; S31875.  
 DR PIR: J02300; J02300.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 DR ATP-binding.  
 KW NP\_BIND 221 228 ATP (BY SIMILARITY).  
 FT SEQUENCE 349 AA: 39722 MW: D5FAE76CD56370F4 CRC64;  
 SQ

Query Match 64.0%; Score 233; DB 1; Length 349;

Best Local Similarity 61.4%; Pred. No. 8 3e-19;  
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLWGEFQVDGSRAGCGQTSNDAAAFALNASSKEALQIRREKPEKYLQFHNINSL 60  
DB 110 TLWGEFQVDGSRAGCGQTSNDAAAFALNASSKEALQIRREKPEKYLQFHNINSL 169  
OY 61 DRIFDKTPEP 70  
DB 170 NRIFQTPPEP 179

RESULT 8  
VAL1\_TYLCM STANDARD; PRT; 359 AA.  
ID P27260;  
AC 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DE 01-JUN-1994 (Rel. 29, Last annotation update)  
GN AL1 protein (C1 protein).

OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).  
OC Viruses; ssDNA viruses; geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10833;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA MEDLINE=92107660; PubMed=1840676;  
RA Khoury-Pour A., Bendahmane M., Matzelt V., Accotto G.P., Crespi S.,  
RA Groenborn B.;  
RT "Tomato yellow leaf curl virus from Sardinia is a  
RT whitefly-transmitted monopartite geminivirus."  
RT Nucleic Acids Res. 19:6763-6769(1991).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X61153; CAA3466.1;  
DR PIR: S22593; S22593.  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1.  
DR PRINTS: PR00227; GEMCOATAL1.  
DR ProDom: PD000736; Gemini\_AL1.  
KW ATP-binding.  
FT NP\_BIND 220  
SQ SEQUENCE 359 AA; 40733 MW; 9717B4A07C93EFA7 CRC64;

Query Match  
Best Local Similarity 63.2%; Score 230; DB 1; Length 359;  
Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

OY 2 TLWGEFQVDGSRAGCGQTSNDAAAFALNASSKEALQIRREKPEKYLQFHNINSL 61  
DB 111 TLWGEFQVDGSRAGCGQTSNDAAAFALNASSKEALQIRREKPEKYLQFHNINSL 170  
OY 62 DRIFDKTPEP 70  
DB 171 KVFQVPAP 179

RESULT 9  
VAL1\_BGMV STANDARD; PRT; 353 AA.  
ID P05175;  
AC 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DE 01-OCT-1996 (Rel. 34, Last annotation update)  
CC

DE AL1 protein (40.2 kDa protein).  
GN AC1.  
OS Bean golden mosaic virus.  
OC Viruses; ssDNA viruses; geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10833;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;  
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene  
RT regulation in geminiviruses."  
RT Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M10070; AAA6318.1;  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1.  
DR PRINTS: PR00227; GEMCOATAL1.  
DR ProDom: PD000736; Gemini\_AL1.  
KW ATP-binding.  
FT NP\_BIND 222  
SQ SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match  
Best Local Similarity 62.6%; Score 228; DB 1; Length 353;  
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLWGEFQVDGSRAGCGQTSNDAAAFALNASSKEALQIRREKPEKYLQFHNINSL 60  
DB 110 TLWGEFQVDGSRAGCGQTSNDAAAFALNASSKEALQIRREKPEKYLQFHNINSL 169  
OY 61 DRIFDKTPEP 70  
DB 170 ERIFKVPPEP 179

RESULT 10  
VAL1\_ABMV STANDARD; PRT; 355 AA.  
ID P21947;  
AC 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DE 01-JUN-1994 (Rel. 29, Last annotation update)  
GN AL1 protein.  
OS Abutilon mosaic virus (isolate West India).  
OC Viruses; ssDNA viruses; geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10816;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA MEDLINE=91020984; PubMed=2219703;  
RA Frischmuth T., Zimmat G., Jeske H.;  
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
RT as well as eukaryotic features."  
RT Virology 178:461-468(1990).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X15983; .: NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; QOCVW1.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding. 221  
 FT NP\_BIND 228  
 SQ SEQUENCE 355 AA: 40257 MW: 16A2CA8A63251E95 CRC64:

Query Match 62.4%; Score 227; DB 1; Length 355;  
 Best Local Similarity 60.0%; Pred. No. 4e-18;  
 Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLVWGEFQVGRSARCGCQTSNDAAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 60  
 DB 110 TIEWGDFQIDGRSARGGQOTANDSYAKALNAGVQSLNLIKEDPKDYVLOHNHNSNL 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 11  
 VAL1\_BCTV STANDARD; PRT; 358 AA.  
 ID VAL1\_BCTV

AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE A11 protein (40.8 kDa protein).  
 OS Beet curly top virus (BCTV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
 OX NCBI\_TaxID=10840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 beet curly top virus."  
 RT EMBO J. 5:1761-1767(1986).  
 RL -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: X04144; .: NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding. 222  
 FT NP\_BIND 229  
 SQ SEQUENCE 358 AA: 40889 MW: 39A5FE3C0B9C333 CRC64:

Query Match 61.5%; Score 224; DB 1; Length 358;  
 Best Local Similarity 57.1%; Pred. No. 8.7e-18;  
 Matches 40; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLVWGEFQVGRSARCGCQTSNDAAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 60  
 DB 110 TIEWGDFQIDGRSARGGQOTANDSYAKALNAGVQSLNLIKEDPKDYVLOHNHNSNL 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 OKIFORPPDP 179

RESULT 12  
 VAL1\_TMOV STANDARD; PRT; 361 AA.  
 ID VAL1\_TMOV

AC Q06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A11 protein.  
 GN A11.  
 OS Tomato mottle virus (isolate Florida) (TMOV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93107858; PubMed=1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 isolated from tomatoes in Florida."  
 RT J. Gen. Virol. 73:3225-3229(1992).  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: L14460; AAC32414.1; .:  
 DR PIR: J01870; J01870.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding. 222  
 FT NP\_BIND 229  
 SQ SEQUENCE 361 AA: 40516 MW: 8138B65CEAC6950 CRC64:

Query Match 61.5%; Score 224; DB 1; Length 361;  
 Best Local Similarity 57.1%; Pred. No. 8.8e-18;  
 Matches 40; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLVWGEFQVGRSARCGCQTSNDAAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 60  
 DB 110 TIEWGDFQIDGRSARGGQOSANDSYAKALNAGVQSLNLIKEDPKDYVLOHNHNSNL 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT; 357 AA.  
 ID VAL1\_TYLCV

AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A11 protein (CI protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 with a single genomic component."  
 RT Virology 185:151-161(1991).  
 RL

-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X15656; CAA33688.1; -  
 DR PIR: D40779; QCCVCL.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1.1.  
 DR PRINTS: PR00227; GEMCOAT1.1.  
 DR ProDom: PD000736; Gemin1\_AL1.1.  
 DR ATP-binding.  
 FM NP-BIND.  
 SO SEQUENCE 357 AA; 40678 MW; 939AB68E1AB2A7 CRC64;

## Query Match

Best Local Similarity 59.1%; Score 215; DB 1; Length 357;  
 Matches 41; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

OY 4 WGFQYDGRSARGCCCTSDNAAAEALNASSKEEALQIREKIPKRYLFOFHNLNSNDR 63  
 Db 111 FGVSDIGRSARGCOOSANDAYAEALNAGSKSEALNITKEKAPDYLLDPHNLNSNDR 170  
 OY 64 F 64  
 Db 171 F 171

## RESULT 14

VAL1\_SLCV STANDARD; PRT; 347 AA.

AC P29048;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE AL1 protein.  
 OS Squash leaf curl virus.  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91082449; PubMed=1984668;  
 RA Lazarowitz S.G., Lazdins I.B.;  
 RT "Infectivity and complete nucleotide sequence of the cloned genomic components of a bipartite squash leaf curl geminivirus with a broad host range phenotype.";  
 RT Virology 180:58-69(1991).  
 RL Virology 180:58-69(1991).

-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: M38183; AAC32410.1; ALT\_INIT.  
 DR PIR: C36785; QCCVSL.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1.1.  
 DR PRINTS: PR00227; GEMCOAT1.1.  
 DR ProDom: PD000736; Gemin1\_AL1.1.  
 DR ATP-binding.  
 FM NP-BIND.  
 SO SEQUENCE 347 AA; 39110 MW; AFDAEBDD122110E CRC64;

## Query Match

Best Local Similarity 34.1%; Score 124; DB 1; Length 347;  
 Matches 25; Conservative 14; Mismatches 23; Indels 4; Gaps 1;

OY 5 GEFQVGRSARGCCCTSDNAAAEALNASSKEEALQIREKIPKRYLFOFHNLNSNDR 64  
 Db 116 GQYKSG-----GSKSNKDDVYHNAVNAGSAGEALDITKADPKFTIVNHNLANVERLF 171  
 OY 65 DKTPPEP 70  
 Db 172 QKPEP 177

## RESULT 15

Y175\_HELPJ STANDARD; PRT; 299 AA.

AC Q92M07;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein JHP0161 precursor.  
 GN JHP0161.  
 OS Helicobacter pylori J99 (Campylobacter pylori J99).  
 CC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OX Helicobacter.  
 OX NCBI\_TaxID=85963;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99120557; PubMed=9923682;  
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
 RA Trust T.J.;  
 RT Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.";  
 RL Nature 397:176-180(1999).  
 CC -1- SIMILARITY: BELONGS TO THE PPTC/PARVULIN FAMILY OF ROTAMASES.  
 CC STRONG. TO C.JEJUNI CBF2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AB001454; AAD05744.1; -  
 DR HSP: Q13526; IPTN.  
 DR InterPro: IPR000297; Rotamase.  
 DR Pfam: PF00639; Rotamase.1.  
 DR PROSITE: PS01096; PPTC\_PPIASE\_1.1.  
 DR PROSITE: PS0198; PPTC\_PPIASE\_2.1.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21  
 FT DOMAIN 154 253  
 SO SEQUENCE 299 AA; 34040 MW; 9C037B1CD1110143 CRC64;

Query Match  
 Best Local Similarity 17.7%; Score 64.5; DB 1; Length 299;  
 Matches 18; Conservative 14; Mismatches 11; Indels 21; Gaps 2;

OY 22 NDAAEALNAS-----SKEEALQIREKIPKRYLFOFHNLNSNDR 62  
 Db 93 NEAKAEKINQPPERKAMEAVKQALVEFNAKQAEVKKIQIPEKMODFY--NANRQ 150  
 OY 63 IPDK 66  
 Db 151 LFVK 154

Thu Oct 10 16:05:31 2002

us-09-289-346a-1.rsp

Page 7

Search completed: October 10, 2002, 02:09:49  
Job time: 1347 sec

---



Thu Oct 10 16:05:30 2002

us-09-289-346a-1.rpr

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:39:22 ; Search time 69.63 Seconds  
(without alignments)  
96.600 Million cell updates/sec

Title: US-09-289-346a-1  
Perfect score: 364  
Sequence: 1 TLVWGEFQVYDGRSARSGCOT.....FOFHNLSNLDRTFDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	364	100.0	352	1	QOCVLI
2	264	72.5	361	1	QOCVPT
3	251	69.0	358	2	S07594
4	244	67.0	362	1	J01887
5	237	65.1	359	2	S39211
6	233	64.0	349	2	J02500
7	230	64.0	349	2	S31875
8	230	63.2	359	2	S22593
9	227	62.4	351	2	J02327
10	227	62.4	355	1	QOCVMI
11	224	61.5	358	1	J01870
12	224	61.5	359	2	S39235
13	224	61.5	385	2	S28360
14	217	59.6	360	2	S59885
15	215	59.1	357	1	QOCVCI
16	142	39.0	131	2	S45059
17	124	34.1	347	1	QOCVSI
18	68	18.7	587	2	JC1419
19	66	18.1	714	2	C95382
20	64.5	17.7	299	2	B71967
21	64.5	17.7	1713	2	A55347
22	63.5	17.4	1610	2	A46227
23	63.5	17.4	1646	2	JH0422
24	63.5	17.4	2161	2	JH0564
25	63.5	17.4	2181	2	A38198
26	62.5	17.2	2203	2	T42742
27	62.5	17.2	447	2	T12544
28	62.5	17.2	642	2	D90558
29	62	17.0	154	2	AD3475

30	62	17.0	181	2	G97976	conserved hypothet
31	62	17.0	295	2	D42452	C1 protein - tobac
32	61.5	16.9	481	2	A70091	probable phospho
33	61	16.8	840	2	T36175	probable large App
34	61	16.8	1502	1	RCBYH1	CYCL/CRP3 transcr
35	60.5	16.6	354	2	A75087	acetyl ornithine d
36	60.5	16.6	1229	2	T48959	kinesin-like prote
37	60	16.5	447	2	S52437	CDP-diacylglycerol
38	59.5	16.3	1033	2	E97700	hypothetical prote
39	59.5	16.3	2137	2	T05244	hypothetical prote
40	59	16.2	160	2	G82060	hypothetical prote
41	59	16.2	316	2	C82085	conserved hypothet
42	59	16.2	338	2	AC3582	conserved hypothet
43	59	16.2	340	2	B97011	iron(III)-bindin
44	59	16.2	397	2	B71078	probably O-sialogl
45	59	16.2	416	2	A82892	probable MADH oxid
						hypothetical prote

# ALIGNMENTS

RESULT 1  
QOCVLI  
All protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
C:Accession: A04170  
R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBL J. 3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus All protein

Query Match 100.0%; Score 364; DB 1; Length 352;  
Best Local Similarity 100.0%; Pred. No. 3.1e-33;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVYDGRSARSGCOTSDAAEALNASSKEEALQIRKIPKYLFOFHNLSNLT 60  
DB 111 TLVWGEFQVYDGRSARSGCOTSDAAEALNASSKEEALQIRKIPKYLFOFHNLSNLT 170  
QY 61 DRIFDKTPPEP 70  
DB 171 DRIFDKTPPEP 180  
RESULT 2  
QOCVPT  
All protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; M01D:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-361 <COU>  
A:References: GB:D00940; M1D:g222458; P1D:N:BAA00782.1; P1D:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus All protein

Query Match 72.5%; Score 264; DB 1; Length 361;  
Best Local Similarity 71.0%; Pred. No. 5,7e-22;  
Matches 49; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60  
DB 110 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60  
QY 61 DRIFDKTPEP 69  
DB 170 DRIFDKTPEP 178

RESULT 3  
S07594

hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)  
C:Species: cassava latent virus  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999

R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.

Nucleic Acids Res. 18, 197-198, 1990

A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava

A:Reference number: S07590; MUID:90174930

A:Accession: S07594

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-358 <MOR>

A:Cross-references: EMBL:X17095; NID:959371; PIDN:CAA34953.1; PID:959376

A:Map position: segment DNAL

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 69.0%; Score 251; DB 2; Length 358;  
Best Local Similarity 64.3%; Pred. No. 1.6e-20;  
Matches 45; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60  
DB 109 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60  
QY 61 DRIFDKTPEP 70  
DB 169 DRIFDKTPEP 178

RESULT 4

AL1 protein - tomato yellow leaf curl virus (strain Australia)  
N:Alternate names: C1 protein  
C:Species: tomato yellow leaf curl virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999

R:Diry, I.B.; Rigden, J.E.; Krahe, L.R.; Mullineaux, P.M.; Rezaiu, M.A.

J. Gen. Virol. 74, 147-151, 1993

A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.

A:Reference number: J01885; MUID:93139778

A:Accession: J01887

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-362 <DRY>

A:Cross-references: GB:S53251

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 67.0%; Score 244; DB 1; Length 362;  
Best Local Similarity 56.5%; Pred. No. 1e-19;  
Matches 48; Conservative 9; Mismatches 12; Indels 16; Gaps 1;

QY 1 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60  
DB 110 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60

DB 110 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 169  
QY 61 DRIFDKTPEP 69  
DB 170 DRIFDKTPEP 178

RESULT 5

gene C1 protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999

R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.

Nucleic Acids Res. 18, 197-198, 1990

A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava

A:Reference number: S39209

A:Accession: S39211

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-359 <MOR>

A:Cross-references: EMBL:225751; NID:9433655; PIDN:CAA81026.1; PID:9433658

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 65.1%; Score 237; DB 2; Length 359;  
Best Local Similarity 62.3%; Pred. No. 6.2e-19;  
Matches 43; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 2 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 61  
DB 111 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 61  
QY 62 DRIFDKTPEP 70  
DB 171 DRIFDKTPEP 179

RESULT 6

replicase - pepper huasteco virus (component A)  
N:Alternate names: ORF AL1 protein  
C:Species: pepper huasteco virus  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999

R:Torres-Pacheco, I.; Garzon-Tirado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante, J. Gen. Virol. 74, 2225-2231, 1993

A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b

A:Reference number: J02299; MUID:94015007

A:Accession: J02300

A:Molecule type: DNA

A:Residues: 1-349 <TOR>

A:Cross-references: GB:X70418; NID:961023; PIDN:CAA49856.1; PID:961025

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 64.0%; Score 233; DB 2; Length 349;  
Best Local Similarity 61.4%; Pred. No. 1.7e-18;  
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60  
DB 110 TLWGEFQVDSARGGCGTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSNL 60  
QY 61 DRIFDKTPEP 70  
DB 170 DRIFDKTPEP 179

RESULT 7

AL1 protein - pepper rizado amarillo virus  
C:Species: pepper rizado amarillo virus

Query Match 64.0%; Score 233; DB 2; Length 349;  
Best Local Similarity 61.4%; Pred. No. 1.7e-18;  
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;







Thu Oct 10 16:05:29 2002

us-09-289-346a-1.rapn

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:44:57 ; Search time 407.82 seconds  
(without alignments)  
47.487 Million cell updates/sec

Title: US-09-289-346a-1  
Perfect score: 364  
Sequence: 1 TLWGEFYDGRSARGGCGT.....FGFHNLSNLDRIFDKPEP 70

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1006125 seqs, 276659714 residues  
Total number of hits satisfying chosen parameters: 1006125

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	20.3	340	6 US-10-219-999-59588	Sequence 59588, A
2	71.5	19.6	945	5 US-09-791-537-132140	Sequence 132140, A
3	68	18.7	587	5 US-09-791-537-33198	Sequence 33198, A
4	67.5	18.5	512	5 US-09-791-537-52432	Sequence 52432, A
5	67.5	18.5	1449	5 US-09-791-537-52442	Sequence 52442, A
6	65	17.9	293	6 US-10-151-336-3	Sequence 3, Appl
7	65	17.9	335	6 US-10-151-336-6	Sequence 6, Appl
8	64.5	17.7	299	5 US-09-791-537-79519	Sequence 79519, A
9	64.5	17.7	1238	5 US-09-791-537-144585	Sequence 144585, A
10	64.5	17.7	1331	5 US-09-791-537-118290	Sequence 118290, A
11	64.5	17.7	1713	5 US-09-791-537-87556	Sequence 87556, A
12	64.5	17.7	1713	5 US-10-171-311-113	Sequence 113, App
13	64.5	17.7	1713	5 US-09-791-537-53064	Sequence 53064, A
14	64.5	17.7	1806	5 US-09-791-537-53064	Sequence 181, App
15	64.5	17.7	3298	7 US-60-365-384-181	Sequence 182, App
16	64.5	17.7	3332	7 US-60-365-384-182	Sequence 182, App
17	63.5	17.4	2181	1 PCT-US02-23913-50	Sequence 50, Appl
18	63.5	17.4	2181	1 PCT-US01-32045-288	Sequence 288, App
19	63.5	17.4	2181	6 US-10-205-823-50	Sequence 50, Appl
20	62.5	17.2	447	5 US-09-991-150-369	Sequence 369, App
21	62.5	17.2	447	5 US-09-989-328-369	Sequence 369, App
22	62.5	17.2	447	6 US-10-103-295-129	Sequence 129, App
23	62.5	17.2	447	6 US-10-121-062-254	Sequence 254, App
24	62.5	17.2	447	6 US-10-173-689-254	Sequence 254, App
25	62.5	17.2	447	6 US-10-173-690-254	Sequence 254, App
26	62.5	17.2	447	6 US-10-173-691-254	Sequence 254, App

27	62.5	17.2	447	6 US-10-173-692-254	Sequence 254, App
28	62.5	17.2	447	6 US-10-173-695-254	Sequence 254, App
29	62.5	17.2	447	6 US-10-173-697-254	Sequence 254, App
30	62.5	17.2	447	6 US-10-173-698-254	Sequence 254, App
31	62.5	17.2	447	6 US-10-173-699-254	Sequence 254, App
32	62.5	17.2	447	6 US-10-173-701-254	Sequence 254, App
33	62.5	17.2	447	6 US-10-173-704-254	Sequence 254, App
34	62.5	17.2	447	6 US-10-173-708-254	Sequence 254, App
35	62.5	17.2	447	6 US-10-174-571-254	Sequence 254, App
36	62.5	17.2	447	6 US-10-174-574-254	Sequence 254, App
37	62.5	17.2	447	6 US-10-174-583-254	Sequence 254, App
38	62.5	17.2	447	6 US-10-173-694-254	Sequence 254, App
39	62.5	17.2	447	6 US-10-173-707-254	Sequence 254, App
40	62.5	17.2	447	6 US-10-174-569-254	Sequence 254, App
41	62.5	17.2	447	6 US-10-174-570-254	Sequence 254, App
42	62.5	17.2	447	6 US-10-174-578-254	Sequence 254, App
43	62.5	17.2	447	6 US-10-174-582-254	Sequence 254, App
44	62.5	17.2	447	6 US-10-174-585-254	Sequence 254, App
45	62.5	17.2	447	6 US-10-174-586-254	Sequence 254, App

ALIGNMENTS

RESULT 1  
US-10-219-999-59588  
Sequence 59588, Application US/10219999  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Edgerton, Michael D  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jiongong  
APPLICANT: Stein, Joshua  
TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-10(52726)C  
CURRENT APPLICATION NUMBER: US/10/219,999  
CURRENT FILING DATE: 2002-08-15  
PRIOR APPLICATION NUMBER: US 60/324,109  
PRIOR FILING DATE: 2001-09-21  
PRIOR APPLICATION NUMBER: US 60/312,544  
PRIOR FILING DATE: 2001-08-15  
NUMBER OF SEQ ID NOS: 63520  
SEQ ID NO 59588  
LENGTH: 340  
TYPE: PRT  
ORGANISM: Zea mays  
US-10-219-999-59588  
Query Match 20.3%; Score 74; DB 6; Length 340;  
Best Local Similarity 26.8%; Pred. No. 1.2;  
Matches 19; Conservative 15; Mismatches 21; Indels 16; Gaps 3;  
OY 15 RGGCTSDAAAEALNASSKEALQ-----TIREKIPEKYLQFHN-----LNS 58  
| | | | | : : : : :  
DB 30 RGGCSKRPPTASATHHCAATKQVRGQKVSDETKMVGEEIPQAWPDSHNLARSPWLNS 89  
OY 59 NLDRIFDKTPPE 69  
| | | | : : : : :  
DB 90 TLSELDKTKQ 100  
RESULT 2  
US-09-791-537-132140  
Sequence 132140, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Danczer, Joseph  
APPLICANT: Danczer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
FILE REFERENCE: 261/210

```

; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132140
; LENGTH: 945
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-791-537-132140

```

```

Query Match
Best Local Similarity 19.6%; Score 71.5; DB 5; Length 945;
Matches 22; Conservative 8; Mismatches 19; Indels 15; Gaps 2;

```

```

OY 10 DGRSARGCQTSNDAAAEALNASSKEPALQIREKIP-----EYLFQPHN 55
Db 145 NGAGVAGGYGYSNNYNA-ALGMLSKERATQREKAKNGRNIVKPKWENLPEFLKDFYN 203
OY 56 LNSN 59
Db 204 IHPN 207

```

```

RESULT 3
US-09-791-537-33198
; Sequence 33198, Application US/09/91537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33198
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Streptococcus sp
US-09-791-537-33198

```

```

Query Match
Best Local Similarity 18.7%; Score 68; DB 5; Length 587;
Matches 18; Conservative 13; Mismatches 21; Indels 8; Gaps 1;
OY 12 RSARGCQTSNDAAAEALNASSKE-----EALQIREKIPKYLFOFHNLSNIDRI 63
Db 114 KSANGKLTALDDADITLQASRVNELSAQNKLQAEAEAAQKALDALNNKNEQIAKL 173

```

```

RESULT 4
US-09-791-537-52432
; Sequence 52432, Application US/09/91537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52432
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-791-537-52432

```

```

Query Match
Best Local Similarity 18.5%; Score 67.5; DB 5; Length 512;
Matches 23; Conservative 11; Mismatches 33; Indels 17; Gaps 3;
OY 2 LWMGEFOYDGRSARGCQTSNDAAAEALNASSKEPALQIREKIPKYL-----L 50
Db 34 LVMASRELEFRACNPGIMDTADGCACKLN--QLEPALKTGSEELRSLYNALAVLYCVHSR 91
OY 51 FQFHNLSNIDRIKPFK-----TPEP 70
Db 92 IQIHNTQEALDKIKIKQEKQHKPEP 115

```

```

RESULT 5
US-09-791-537-52442
; Sequence 52442, Application US/09/91537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52442
; LENGTH: 1449
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-791-537-52442

```

```

Query Match
Best Local Similarity 18.5%; Score 67.5; DB 5; Length 1449;
Matches 23; Conservative 11; Mismatches 33; Indels 17; Gaps 3;
OY 2 LWMGEFOYDGRSARGCQTSNDAAAEALNASSKEPALQIREKIPKYL-----L 50
Db 34 LVMASRELEFRACNPGIMDTADGCACKLN--QLEPALKTGSEELRSLYNALAVLYCVHSR 91
OY 51 FQFHNLSNIDRIKPFK-----TPEP 70
Db 92 IQIHNTQEALDKIKIKQEKQHKPEP 115

```

```

RESULT 6
US-10-151-336-3
; Sequence 3, Application US/10151336
; GENERAL INFORMATION:
; APPLICANT: Mason, Hugh
; APPLICANT: Palmer, Kenneth
; APPLICANT: Hefteron, Kathleen
; APPLICANT: Mor, Tsafrit
; TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants
; FILE REFERENCE: 4868/84453
; CURRENT APPLICATION NUMBER: US/10/151,336
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/414,276
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Bean yellow dwarf virus
US-10-151-336-3

```

Query Match 17.9%; Score 65; DB 6; Length 293;  
 Best Local Similarity 33.3%; Pred. No. 13;  
 Matches 13; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

OY 31 ASSKEALQIIRKIEPKYLFQFHNLNSNDRIFDKTPE 69  
 DB 133 ATSKEEYLDIKKEPHEMATKLOMLEYSANKLFPPOPE 171

## RESULT 7

US-10-151-336-6  
 ; Sequence 6. Application US/10151336  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mason, Hugh  
 ; APPLICANT: Palmer, Kenneth  
 ; APPLICANT: Hefteron, Kathleen  
 ; APPLICANT: Mor, Tsafir  
 ; APPLICANT: Airlzen, Charles  
 ; TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants  
 ; FILE REFERENCE: 4868/84453  
 ; CURRENT APPLICATION NUMBER: US/10/151,336  
 ; PRIOR FILING DATE: 2002-05-20  
 ; PRIOR APPLICATION NUMBER: US/09/414,276  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6  
 ; LENGTH: 335  
 ; TYPE: PRT  
 ; ORGANISM: Bean yellow dwarf virus  
 ; US-10-151-336-6

Query Match 17.9%; Score 65; DB 6; Length 335;  
 Best Local Similarity 33.3%; Pred. No. 15;  
 Matches 13; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

OY 31 ASSKEALQIIRKIEPKYLFQFHNLNSNDRIFDKTPE 69  
 DB 133 ATSKEEYLDIKKEPHEMATKLOMLEYSANKLFPPOPE 171

## RESULT 8

US-09-791-537-79519  
 ; Sequence 79519, Application US/09791537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bionomix, Inc.  
 ; APPLICANT: Debe, Derek  
 ; APPLICANT: Danzer, Joseph  
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
 ; FILE REFERENCE: 261/210  
 ; CURRENT APPLICATION NUMBER: US/09/791,537  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 153055  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 79519  
 ; LENGTH: 299  
 ; TYPE: PRT  
 ; ORGANISM: Helicobacter pylori  
 ; US-09-791-537-79519

Query Match 17.7%; Score 64.5; DB 5; Length 299;  
 Best Local Similarity 28.1%; Pred. No. 15;  
 Matches 18; Conservative 14; Mismatches 11; Indels 21; Gaps 2;

OY 22 NDAAAEALNLS-----SKERALQIIRKIEPKYLFQFHNLNSNDR 62  
 DB 93 NEAKAKELNQTDEPKAMMEAVKQALVEFWAKQAEEVKIQLPEKEMQDFY--NANKDO 150  
 OY 63 IFDK 66  
 :| |

DB 151 LEVK 154

## RESULT 9

US-09-791-537-144585  
 ; Sequence 144585, Application US/09791537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bionomix, Inc.  
 ; APPLICANT: Debe, Derek  
 ; APPLICANT: Danzer, Joseph  
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
 ; FILE REFERENCE: 261/210  
 ; CURRENT APPLICATION NUMBER: US/09/791,537  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 153055  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 144585  
 ; LENGTH: 1238  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-791-537-144585

Query Match 17.7%; Score 64.5; DB 5; Length 1238;  
 Best Local Similarity 34.4%; Pred. No. 93;  
 Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

OY 9 VDGRSARGGCGTSDNDAAEALN--ASSKEALQIIRKIEPKYLFQFHNLNSNDRIFD 65  
 DB 553 VDAVAVENILNMIKAEDANRAASASALSQIYIKEDLPK----AKTLSSNSDKLN 608

OY 66 K 66  
 DB 609 E 609

## RESULT 10

US-09-791-537-118290  
 ; Sequence 118290, Application US/09791537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bionomix, Inc.  
 ; APPLICANT: Debe, Derek  
 ; APPLICANT: Danzer, Joseph  
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
 ; FILE REFERENCE: 261/210  
 ; CURRENT APPLICATION NUMBER: US/09/791,537  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 153055  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 118290  
 ; LENGTH: 1331  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-791-537-118290

Query Match 17.7%; Score 64.5; DB 5; Length 1331;  
 Best Local Similarity 34.4%; Pred. No. 1e+02;  
 Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

OY 9 VDGRSARGGCGTSDNDAAEALN--ASSKEALQIIRKIEPKYLFQFHNLNSNDRIFD 65  
 DB 646 VDAVAVENILNMIKAEDANRAASASALSQIYIKEDLPK----AKTLSSNSDKLN 701  
 OY 66 K 66  
 DB 702 E 702

RESULT 11  
 US-09-643-992A-2

```

; Sequence 2, Application US/09643992A
; GENERAL INFORMATION:
; APPLICANT: Oriental Yeast Co., Ltd.
; APPLICANT: NAITOU Toshikuni
; TITLE OF INVENTION: Method for neuragenesis
; FILE REFERENCE: EP/0-21-805
; CURRENT APPLICATION NUMBER: US/09/643,992A
; CURRENT FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 1713
; TYPE: PRT
; ORGANISM: Human
US-09-643-992A-2

```

```

Query Match          17.7%; Score 64.5; DB 5; Length 1713;
Best Local Similarity 34.4%; Pred. No. 1.4e+02;
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```

```

OY 9 VDGSRAGGCGTSDAAAEALN--ASSKEPALQ-IIRKIPKYLQFPHNLSNLDRIFD 65
Db 553 VDAATAYENTILNAIKAEADANRAASASASALQTVIKEDLPK-----AKTSSNSDKLLN 608
OY 66 K 66
Db 609 E 609

```

```

RESULT 12
US-09-791-537-87556
; Sequence 87556, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87556
; LENGTH: 1713
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-87556

```

```

Query Match          17.7%; Score 64.5; DB 5; Length 1713;
Best Local Similarity 34.4%; Pred. No. 1.4e+02;
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;
OY 9 VDGSRAGGCGTSDAAAEALN--ASSKEPALQ-IIRKIPKYLQFPHNLSNLDRIFD 65
Db 553 VDAATAYENTILNAIKAEADANRAASASASALQTVIKEDLPK-----AKTSSNSDKLLN 608
OY 66 K 66
Db 609 E 609

```

```

RESULT 13
US-10-171-311-113
; Sequence 113, Application US/10171311
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatz, Karen

```

```

; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersth, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 1713
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-113

```

```

Query Match          17.7%; Score 64.5; DB 6; Length 1713;
Best Local Similarity 34.4%; Pred. No. 1.4e+02;
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```

```

OY 9 VDGSRAGGCGTSDAAAEALN--ASSKEPALQ-IIRKIPKYLQFPHNLSNLDRIFD 65
Db 553 VDAATAYENTILNAIKAEADANRAASASASALQTVIKEDLPK-----AKTSSNSDKLLN 608
OY 66 K 66
Db 609 E 609

```

```

RESULT 14
US-09-791-537-53064
; Sequence 53064, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53064
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53064

```

```

Query Match          17.7%; Score 64.5; DB 5; Length 1806;
Best Local Similarity 34.4%; Pred. No. 1.5e+02;
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```

```

OY 9 VDGSRAGGCGTSDAAAEALN--ASSKEPALQ-IIRKIPKYLQFPHNLSNLDRIFD 65
Db 646 VDAATAYENTILNAIKAEADANRAASASASALQTVIKEDLPK-----AKTSSNSDKLLN 701
OY 66 K 66
Db 702 E 702

```

```

RESULT 15
US-60-365-384-181
; Sequence 181, Application US/60365384

```

## GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom  
 APPLICANT: Wang, Jian-rui  
 APPLICANT: Wang, Zhiwei  
 APPLICANT: Zhang, Jie  
 APPLICANT: Zhou, Ping  
 APPLICANT: Wehrman, Tom  
 APPLICANT: Wang, Jian-Rui  
 APPLICANT: Ghosh, Malabika  
 APPLICANT: Zhao, Qing A.  
 APPLICANT: Asundi, Vinod  
 APPLICANT: Ren, Feiyao  
 APPLICANT: Xue, Aidong J.  
 APPLICANT: Ma, Yungqing  
 APPLICANT: Wang, Dunrui  
 APPLICANT: Weng, Gezhi  
 TITLE OF INVENTION: Novel Nucleic Acids and  
 TITLE OF INVENTION: Polypeptides  
 FILE REFERENCE: 814  
 CURRENT APPLICATION NUMBER: US/60/365,384  
 CURRENT FILING DATE: 2002-03-14  
 NUMBER OF SEQ ID NOS: 666  
 SOFTWARE: pf-fl-genes Version 6.0  
 SEQ ID NO 181  
 LENGTH: 3298  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-60-365-384-181

## Query Match

17.7%; Score 64.5; DB 7; Length 3298;

Best Local Similarity 34.4%; Pred. No. 3.2e+02; Indels 7; Gaps 3;

Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

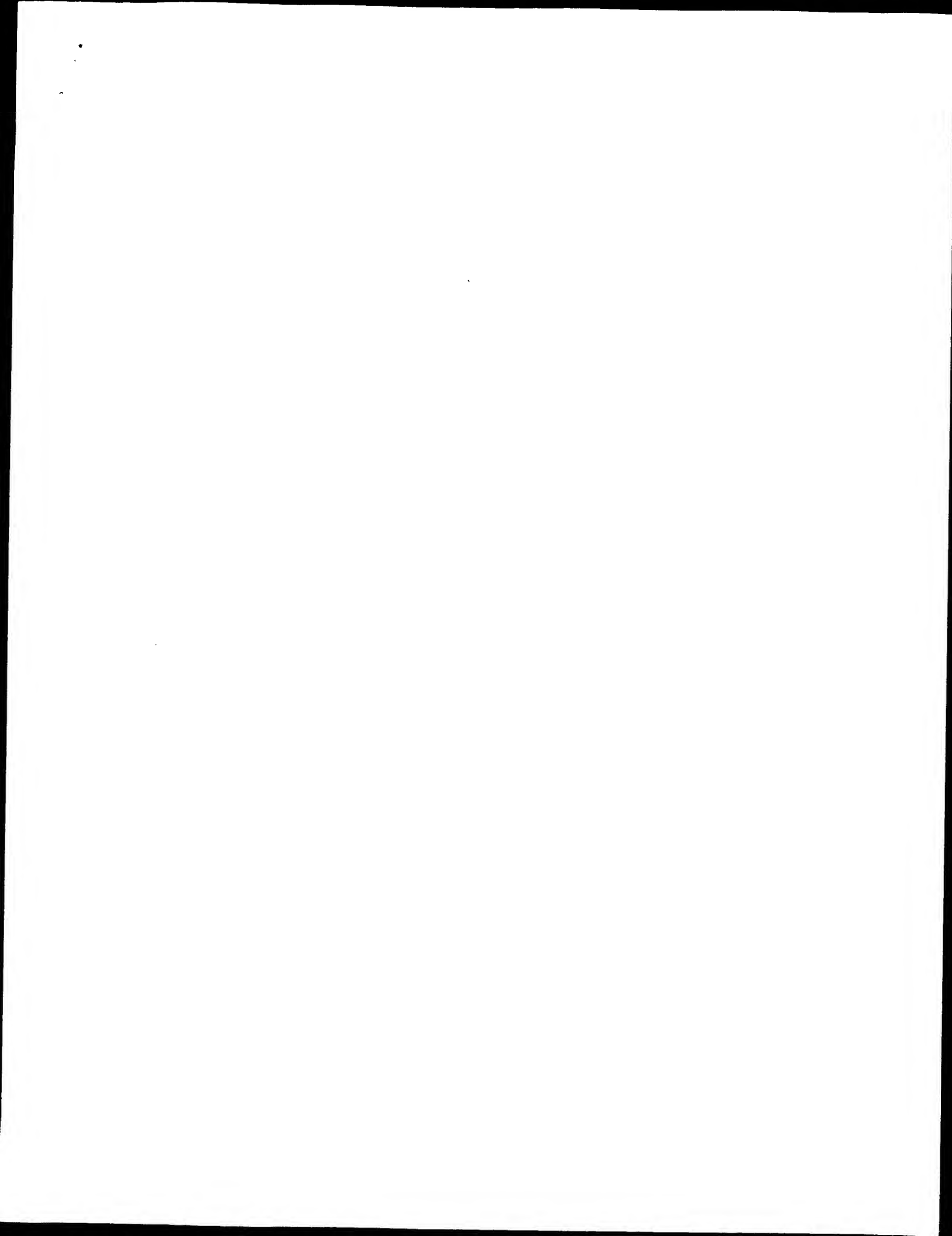
QY 9 VDGSRAGCGCOTSDAAAEALN--ASSKEEALQ-IIRKIPKYLFOFHNLSNLDRIED 65

DB 2138 VDAATAYENTLNATKAAEDANRAASASSESALQTVTKEDLPK---AKTLSSNSDKLLN 2193

OY 66 K 66

DB 2194 E 2194

Search completed: October 10, 2002, 02:09:00  
 Job time: 1443 sec





Thu Oct 10 16:05:43 2002

us-09-289-346a-2.ispt

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds  
(Without alignments)  
102.295 Million cell updates/sec

Title: us-09-289-346a-2  
Perfect score: 362  
Sequence: 1 TLVWGEFOVDGAAAGGCCT.....FOFHNLNSLDRIFDKTPEP 70

Scoring table: BL0SUM62  
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

- 1: SP\_ARCHAEA:\*
- 2: SP\_BACTERIA:\*
- 3: SP\_FUNGI:\*
- 4: SP\_HUMAN:\*
- 5: SP\_INVERTEBRATE:\*
- 6: SP\_MAMMAL:\*
- 7: SP\_MHC:\*
- 8: SP\_ORGANELLE:\*
- 9: SP\_PHAGE:\*
- 10: SP\_PLANT:\*
- 11: SP\_RODENT:\*
- 12: SP\_VIRUS:\*
- 13: SP\_VERTEBRATE:\*
- 14: SP\_UNCLASSIFIED:\*
- 15: SP\_VIRUS:\*
- 16: SP\_BACTERIAP:\*
- 17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	82.9	351	12 Q91R10	Q91R10 tomato seve
2	298	82.3	352	12 Q9E000	Q9E000 tomato rugo
3	281	77.6	226	12 Q09727	Q09727 leonurus mo
4	281	77.6	226	12 Q9WHF6	Q9WHF6 tomato mild
5	279	77.1	361	12 Q67574	Q67574 bean golden
6	277	76.5	225	12 Q9QDPI	Q9QDPI compesa gold
7	277	76.5	314	12 Q9ELT8	Q9ELT8 sweet potat
8	274	75.7	364	12 Q9Q555	Q9Q555 sweet potat
9	274	75.7	185	12 Q9B975	Q9B975 sida golden
10	266	73.5	149	12 Q9B975	Q9B975 sida golden
11	266	73.5	233	12 Q9YLA4	Q9YLA4 macropitiliu
12	260	71.8	234	12 Q9180	Q9180 geminivirid
13	257	71.0	190	12 Q92089	Q92089 tobacco lea
14	257	71.0	190	12 Q92084	Q92084 tobacco lea
15	256	70.7	190	12 Q9W827	Q9W827 tobacco lea
16	256	70.7	208	12 Q920C4	Q920C4 tobacco lea

17	256	70.7	359	12 Q91M88	Q91M88 tobacco lea
18	254	70.2	208	12 Q920C0	Q920C0 tobacco lea
19	254	70.2	208	12 Q920B8	Q920B8 tobacco lea
20	252	69.6	208	12 Q920H6	Q920H6 tobacco lea
21	251	69.3	203	12 Q92083	Q92083 tobacco lea
22	251	69.3	363	12 Q92577	Q92577 cotton leaf
23	246	68.0	359	12 Q91B86	Q91B86 ageratum ye
24	246	68.0	360	12 Q90X10	Q90X10 ageratum ye
25	244	67.4	208	12 Q920A0	Q920A0 tobacco lea
26	244	67.4	363	12 Q92705	Q92705 cotton leaf
27	244	67.4	363	12 Q92719	Q92719 cotton leaf
28	243	67.1	190	12 Q92086	Q92086 tobacco lea
29	243	67.1	208	12 Q920C6	Q920C6 tobacco lea
30	242	66.9	349	12 Q88888	Q88888 tomato pseu
31	242	66.9	362	12 Q56816	Q56816 chayote mos
32	241	66.6	359	12 Q9Y2V4	Q9Y2V4 tomato yell
33	241	66.6	359	12 Q9Y2V2	Q9Y2V2 tomato yell
34	241	66.6	359	12 Q9YU07	Q9YU07 tomato yell
35	241	66.6	359	12 Q9YU27	Q9YU27 tomato yell
36	240	66.3	348	12 Q91W5	Q91W5 macropitiliu
37	240	66.3	353	12 Q72692	Q72692 beet curly
38	240	66.3	354	12 Q65438	Q65438 beet curly
39	240	66.3	359	12 Q88942	Q88942 tomato yell
40	239	66.0	190	12 Q920A7	Q920A7 tobacco lea
41	239	66.0	362	12 Q920A7	Q920A7 tobacco lea
42	238	65.7	307	12 Q91N48	Q91N48 okra enatio
43	238	65.7	361	12 Q91E77	Q91E77 cotton leaf
44	237	65.5	231	12 Q72723	Q72723 cotton leaf
45	237	65.5	354	12 Q91N42	Q91N42 african tom

ALIGNMENTS

RESULT 1  
ID Q91R10 PRELIMINARY: PRT: 351 AA.  
AC Q91R10;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE REP PROTEIN.  
GN AC1.  
OS Tomato severe rugose virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=158463;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MINAS GERARS:  
RA Rezende W.L., Goulart L.R., Patreira K.S., Figueiredo J.E.F.;  
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting  
RT begomovirus, Tomato severe rugose virus, in Brazil."  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY029750; AAK50357.1; -  
SQ SEQUENCE 351 AA: 40122 MW: 87F937A4F873B6CF CRC64;

Query Match 82.9%; Score 300; DB 12; Length 351;  
Best local similarity 78.6%; Pred. No. 3.9e-25;  
Matches 55; Conservative 8; Mismatches 7; Indels 0; Gaps 0;  
OY 1 TLVWGEFOVDGAAAGGCCTSDNDAAEALNASSKEEALQITREKIPERYLFQFHNLNSL 60  
DB 111 TLVWGEFOVDGSAAGGCCTANDAAAEALNASSKEEALQITREKIPERYLFQFHNLNSL 170  
OY 61 DRIFDKTPEP 70  
DB 171 DRIFDKTPEP 180  
RESULT 2  
ID Q9E000 PRELIMINARY: PRT: 352 AA.

AC 09E000;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN.  
 GN AC1.  
 OS Tomato rugose mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=134599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,  
 RA Zambolim E.M., Zerbini F.M.;  
 RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
 RT (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,  
 RT Minas Gerais, Brazil.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF291705; AAG15546.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; GEMCOATL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Geminl\_AL1; 1.  
 SQ SEQUENCE 352 AA; 40012 MW; 47CD5838E2AD613 CRC64;

Query Match 82.3%; Score 298; DB 12; Length 352;  
 Best Local Similarity 78.6%; Pred. No. 6,3e-25;  
 Matches 55; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGAAAGCOTSDNAAEALNASSKEEALQIREKPEKYLFOFHNLSNL 60  
 DB 111 TLVWGEFVIDGRSARGCOTANDAAEALNAPSKNLQIQIREKPEKFLFOFHNLSNL 170  
 OY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 3  
 ID 009727 PRELIMINARY; PRT; 226 AA.  
 AC 009727;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REP. PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonurus mosaica virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=58177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-LEWY-BRAZIL 1;  
 RA Faria J.C., Maxwell D.P.;  
 RT "Variability in geminivirus associated with Phaseolus vulgaris in  
 RT Brazil.";  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U92532; AAB51157.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; GEMCOATL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 226  
 FT 226  
 SQ SEQUENCE 226 AA; 25617 MW; 73CD86E766083FC5 CRC64;

Query Match 77.6%; Score 281; DB 12; Length 226;  
 Best Local Similarity 77.1%; Pred. No. 3e-23;  
 Matches 54; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGAAAGCOTSDNAAEALNASSKEEALQIREKPEKYLFOFHNLSNL 60  
 DB 111 TLVWGEFVDGSSRGGOOTVNDAAEALNAPDKRTALQIQIREKPEKFLFOFHNLSNL 170

OY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 4  
 ID 09WHF6 PRELIMINARY; PRT; 226 AA.  
 AC 09WHF6;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS tomato mild mottle geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
 OX NCBI\_TaxID=92943;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-HN96-HSKW;  
 RA Nakha M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
 RA Maxwell D.P.;  
 RT "Molecular characterization and DNA-based detection methods for  
 RT vegetable-infecting geminiviruses in Central America.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF131071; AAD34471.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; GEMCOATL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 226  
 FT 226  
 SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 77.6%; Score 281; DB 12; Length 226;  
 Best Local Similarity 72.9%; Pred. No. 3e-23;  
 Matches 51; Conservative 13; Mismatches 6; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGAAAGCOTSDNAAEALNASSKEEALQIREKPEKYLFOFHNLSNL 60  
 DB 111 TLVWGEFVIDGRSARGCOTANDAAEALNASSKEEAMRITREKPEKFLFOFHNLSNL 170  
 OY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 5  
 ID 067574 PRELIMINARY; PRT; 361 AA.  
 AC 067574;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN ALL.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 RA Maxwell D.P., Russell D.R.;  
 RT "Cloning of the complete DNA genomes of four bean-infecting  
 RT geminiviruses and determining their infectivity by electric discharge  
 RT particle acceleration.";  
 RL Phytopathology 81:980-985(1991).  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 RA Morales F.J., Maxwell D.P.;  
 RT "Differentiation of bean-infecting geminiviruses by nucleic acid

hybridization probes and aspects of bean golden mosaic in Brazil.",  
 Plant Dis. 75:336-342(1991).  
 (3)  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.:  
 "Genetic diversity in geminiviruses causing bean golden mosaic  
 disease: The nucleotide sequence of the infectious cloned DNA  
 components of a Brazilian isolate of bean golden mosaic geminivirus."  
 RT Submitted (May-1992) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: M88696; AAA46312.1; "  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.1%; Score 279; DB 12; Length 361;  
 Best Local Similarity 76.1%; Pred. No. 8.4e-23;  
 Matches 51; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 4 WGEFVYDGAAGGCGTSDNDAAEALNNSKEEALQIIRKIPKYLFOFHNLSNLDRI 63  
 DB 113 WGHQVYDGRSARGGQOTANDAAEALNNSKEEAMQIIRKIPKYLFOFHNLSNLDRI 172  
 QY 64 FDKTPEP 70  
 DB 173 FTKAPDP 179

RESULT 6  
 Q90DB1 PRELIMINARY; PRT; 225 AA.  
 ID Q90DB1  
 AC 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=69263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGW-BR;  
 RA Faria J.C.;  
 RT "partial nucleotide sequence of cowpea golden mosaic geminivirus from  
 Brazil."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF186708; AAF06318.1; "  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 225  
 SO SEQUENCE 225 AA; 25766 MW; 1089CB6BD8D15B5D CRC64;

Query Match 76.5%; Score 277; DB 12; Length 225;  
 Best Local Similarity 76.1%; Pred. No. 8.2e-23;  
 Matches 51; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 4 WGEFVYDGAAGGCGTSDNDAAEALNNSKEEALQIIRKIPKYLFOFHNLSNLDRI 63  
 DB 113 WGHQVYDGRSARGGQOTANDAAEALNNSKEEAMQIIRKIPKYLFOFHNLSNLDRI 172  
 QY 64 FDKTPEP 70  
 DB 173 FTKAPDP 179

RESULT 7  
 Q9ELT8

Q9ELT8 PRELIMINARY; PRT; 314 AA.  
 AC Q9ELT8  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States."  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 DR EMBL: AF288227; AAC01006.1; "  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 76.5%; Score 277; DB 12; Length 314;  
 Best Local Similarity 79.4%; Pred. No. 1.2e-22;  
 Matches 54; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFVYDGAAGGCGTSDNDAAEALNNSKEEALQIIRKIPKYLFOFHNLSNLDRI 60  
 DB 110 TLWGEFVYDGRSARGGQOTANDAAEALNNSKEEALQIIRKIPKYLFOFHNLSNLDRI 169  
 QY 61 DRIFDKTP 68  
 DB 170 DRIFSPPP 177

RESULT 8  
 Q90S55 PRELIMINARY; PRT; 364 AA.  
 ID Q90S55  
 AC Q90S55  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States."  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl  
 virus: nucleotide sequence and phylogenetic relationships."  
 RL Online Publication.  
 DR EMBL: AF104036; AAD47173.1; "  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 364 AA; 40680 MW; 5F79752431A09D06 CRC64;

Query Match  
Best Local Similarity 76.5%; Score 277; DB 12; Length 364;  
Matches 54; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDDGAAAGCQTSNDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 60  
DB 110 TITMGVFOIDGSRARGGQOTANDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 169  
QY 61 DRIFDKTPP 68  
DB 170 DRIFSPPP 177

RESULT 9

Q98693 PRELIMINARY; PRT; 185 AA.

AC Q98693; 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DE REP PROTEIN (FRAGMENT).  
GN AC1.  
OS sida golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51034;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA;  
RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
RT "Genetic diversity among geminiviruses associated with the weed  
RT species Sida spp. Macroptilium lathyroides, and Wissadula amplissima  
RL Plant Dis. 81:1251-1258(1997).  
DR EMBL; 067926; AAB97865.1; -;  
DR InterPro; IPR001191; Geminl\_AL1.  
DR Pfam; PF00799; Geminl\_AL1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR Prodom; PD000736; Geminl\_AL1.  
FT NON\_TER 1  
FT 185  
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5FE1 CRC64;

Query Match  
Best Local Similarity 75.7%; Score 274; DB 12; Length 185;  
Matches 49; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDDGAAAGCQTSNDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 60  
DB 89 TIEGCVFOIDGSRARGGQOTANDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 148  
QY 61 DRIFDKTPP 70  
DB 149 DRIFSKPP 158

RESULT 10

P88975 PRELIMINARY; PRT; 149 AA.

AC P88975; 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
GN AC1.  
OS Macroptilium golden mosaic geminivirus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA;  
RA Roye M.E.;  
RL Mejis (1996), Biochemistry, University of the West Indies, Jamaica.

DR EMBL; 075278; AAB36919.1; -;  
DR InterPro; IPR001191; Geminl\_AL1.  
DR Pfam; PF00799; Geminl\_AL1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR Prodom; PD000736; Geminl\_AL1.  
FT NON\_TER 1  
FT 149  
SQ SEQUENCE 149 AA; 16785 MW; E4CF5EDD4C9CD508 CRC64;

Query Match  
Best Local Similarity 73.5%; Score 266; DB 12; Length 149;  
Matches 48; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDDGAAAGCQTSNDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 60  
DB 52 TIEGCVFOIDGSRARGGQOTANDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 111  
QY 61 DRIFDKTPP 70  
DB 112 DRIFKDP 121

RESULT 11

Q9YL4 PRELIMINARY; PRT; 233 AA.

AC Q9YL4; 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
GN RRP.  
OS Macroptilium golden mosaic geminivirus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA STRAIN 1;  
RA Roye M.E.;  
RT "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses  
RT from Jamaica.";  
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA STRAIN 1;  
RA Roye M.E., McLaughlin W.A., Maxwell D.P.;  
RT "Molecular characterization of two distinct geminiviruses infecting M.  
RT lathyroides from Jamaica.";  
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF098940; AAD17850.1; -;  
DR InterPro; IPR001191; Geminl\_AL1.  
DR Pfam; PF00799; Geminl\_AL1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR Prodom; PD000736; Geminl\_AL1.  
FT NON\_TER 233  
FT 233  
SQ SEQUENCE 233 AA; 26356 MW; AA490AFAD2166A02 CRC64;

Query Match  
Best Local Similarity 73.5%; Score 266; DB 12; Length 233;  
Matches 48; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDDGAAAGCQTSNDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 60  
DB 110 TIEGCVFOIDGSRARGGQOTANDAAAEALNASSKEEALQITREKIPKYLFOFHNLSNL 169  
QY 61 DRIFDKTPP 70  
DB 170 DRIFKDP 179

RESULT 12

O39180

```

RESULT 15
09W827
ID 09W827 PRELIMINARY; PRT: 190 AA.
AC 09W827;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORAL-1, PARTIAL AND
DE COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_Taxid=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GORAL;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT J. Plant Res. 110:247-257(1997).
RL EMBL: AB001303; BAA34010.1; -.
DR InterPro: IPR001191; Geminl.AL1.
DR Pfam: PF00789; Geminl.AL1.
DR PRINTS: PF00227; GEMCOATL1
DR ProDom: PD000736; Geminl.AL1.

```



Thu Oct 10 16:05:41 2002

us-09-289-346a-2.rpr

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:42 ; Search time 69.63 Seconds  
(without alignments)  
96,600 Million cell updates/sec

Title: US-09-289-346a-2  
Perfect score: 362  
Sequence: 1 TLVWGEFQVDGAAAGGCGT.....PQFHNLSNDRIDFKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	349	96.4	352	1	QOCVLI
2	249	68.8	361	1	QOCVPT
3	236	65.2	358	2	S07594
4	229	63.3	362	1	J01887
5	222	61.3	359	2	S39211
6	218	60.2	349	2	J02300
7	215	59.4	349	2	S31875
8	212	58.6	351	2	J02327
9	212	58.6	355	1	QOCVW1
10	209	57.7	358	1	J01870
11	209	57.7	359	2	S39235
12	209	57.7	385	2	S28360
13	202	55.8	360	2	S59885
14	200	55.2	357	2	QOCVCI
15	142	39.2	131	1	S45059
16	124	34.3	347	1	QOCVSI
17	124	34.3	347	1	A55347
18	68.5	18.9	1713	2	A46227
19	65.5	18.1	1610	2	JH0422
20	65.5	18.1	1646	2	JH0564
21	65.5	18.1	2161	2	A38198
22	65.5	18.1	2181	2	T42742
23	65.5	18.1	2203	2	B69502
24	65	18.0	371	2	B71967
25	64.5	17.8	299	2	DEBSGF
26	64.5	17.8	335	1	A47162
27	64.5	17.8	557	2	T45290
28	63.5	17.5	392	2	B71078
29	63	17.4	397	2	B71078

30	62.5	17.3	419	2	S73846
31	62.5	17.3	1033	2	E97700
32	62	17.1	154	2	AD3475
33	62	17.1	181	2	G97976
34	62	17.1	295	2	D42452
35	62	17.1	308	2	C84072
36	62	17.1	1502	1	RGBYH1
37	61.5	17.0	481	2	A70091
38	61	16.9	338	2	AG3582
39	61	16.9	392	2	T44362
40	61	16.9	587	2	JC1419
41	60.5	16.7	136	2	T22240
42	60.5	16.7	201	2	A81380
43	60.5	16.7	652	2	E89841
44	60	16.6	384	1	M2ML41
45	60	16.6	447	2	S52437

#### ALIGNMENTS

RESULT 1

QOCVLI  
ALI protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 96.4% Score 349; DB 1; Length 352;  
Best Local Similarity 95.7% Pred. No. 1.5e-30;  
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGAAAGGCGTSDNAAPALNASSKEPALQIRREKIPKYLQFPHNLSNL 60  
||||||| : |||||||  
DB 111 TLVWGEFQVDGRSARGCGTSDNAAPALNASSKEPALQIRREKIPKYLQFPHNLSNL 170

OY 61 DRIFDKTPPEP 70  
|||||||  
DB 171 DRIFDKTPPEP 180

RESULT 2

QOCVPT  
ALI protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C:Accession: J00364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: J00362; MUID:91311403

A:Accession: J00364

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1361 <COU>

A:Cross-references: GB:D00940; NID:9222458; PIDN:BAA00782.1; PID:9222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 68.8%; Score 249; DB 1; Length 361;  
 Best Local Similarity 66.7%; Pred. No. 1.4e-19;  
 Matches 46; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

OY 1 TLWGEFQVDCAAAAGCGCOTSDNDAAEALNASSKEALQITREKIPKRYLFQPHNLNSNL 60  
 Db 110 TLEMGEFQIDGRSARGGOQSDANDAYAKALNAGSKSEALNVLRELAPKDYVLOPHNLNSNL 169  
 OY 61 DRIFDKTPEP 69  
 Db 170 DRIFDKTPEP 178

RESULT 3  
 S07594  
 hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)  
 C:Species: cassava latent virus  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
 C:Accession: S07594

R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.  
 Nucleic Acids Res. 18, 197-198, 1990  
 A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava

A:Reference number: S07590; M0ID:90174930  
 A:Accession: S07594  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <MOR>

A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376  
 C:Genetics:  
 A:Map position: segment DNA1  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 65.2%; Score 236; DB 2; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 3.6e-18;  
 Matches 42; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLWGEFQVDCAAAAGCGCOTSDNDAAEALNASSKEALQITREKIPKRYLFQPHNLNSNL 60  
 Db 109 TLEMGEFQIDGRSARGGOQSDANDAYAKALNAGSKSEALNVLRELAPKDYVLOPHNLNSNL 168  
 OY 61 DRIFDKTPEP 70  
 Db 169 DRIFDKTPEP 178

RESULT 4  
 J01887  
 AL1 protein - tomato yellow leaf curl virus (strain Australia)

N:Alternate names: CI protein  
 C:Species: tomato yellow leaf curl virus

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01887

R:Dry, I.B.; Rigden, J.E.; Kirke, L.R.; Mullineaux, P.M.; Rezaian, M.A.  
 J. Gen. Virol. 74, 147-151, 1993

A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.  
 A:Reference number: J01885; M0ID:93139778  
 A:Accession: J01887

A:Status: translation not shown  
 A:Molecule type: DNA

A:Residues: 1-362 <DRY>  
 A:Cross-references: GB:S53251

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 63.3%; Score 229; DB 1; Length 362;  
 Best Local Similarity 52.9%; Pred. No. 2.1e-17;  
 Matches 45; Conservative 10; Mismatches 14; Indels 16; Gaps 1;

OY 1 TLWGEFQVDCAAAAGCGCOTSDNDAAEALNASSKEALQITREKIPKRYLFQPHNLNSNL 60  
 Db 110 TLEMGEFQIDGRSARGGOQSDANDAYAKALNAGSKSEALNVLRELAPKDYVLOPHNLNSNL 169

Db 110 TLEMGEFQIDGRSARGGOQSDANDAYAKALNAGSKSEALNVLRELAPKDYVLOPHNLNSNL 169  
 OY 61 DRIFDKTPEP 69  
 Db 170 DRIFDKTPEP 178

RESULT 5  
 S39211

gene C1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
 C:Accession: S39211

R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.  
 Nucleic Acids Res. 18, 197-198, 1990

A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava  
 A:Reference number: S07590; M0ID:90174930  
 A:Accession: S07594

A:Status: Preliminary  
 A:Molecule type: DNA

A:Residues: 1-359 <MOR>  
 A:Cross-references: EMBL:X25751; NID:g433655; PIDN:CAA81026.1; PID:g433658

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.3%; Score 222; DB 2; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 1.2e-16;  
 Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

OY 2 TLWGEFQVDCAAAAGCGCOTSDNDAAEALNASSKEALQITREKIPKRYLFQPHNLNSNL 61  
 Db 111 TLEMGEFQIDGRSARGGOQSDANDAYAKALNAGSKSEALNVLRELAPKDYVLOPHNLNSNL 170  
 OY 62 DRIFDKTPEP 70  
 Db 171 DRIFDKTPEP 178

RESULT 6  
 J02300

replicase - pepper huasteco virus (component A)  
 N:Alternate names: ORF AL1 protein

C:Species: pepper huasteco virus  
 C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999  
 C:Accession: J02300

R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,  
 J. Gen. Virol. 74, 2225-2231, 1993

A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b  
 A:Reference number: J02299; M0ID:94015007  
 A:Accession: J02300

A:Status: DNA  
 A:Molecule type: DNA

A:Residues: 1-349 <TOR>  
 A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA9856.1; PID:g61025

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.2%; Score 218; DB 2; Length 349;  
 Best Local Similarity 57.1%; Pred. No. 3.3e-16;  
 Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLWGEFQVDCAAAAGCGCOTSDNDAAEALNASSKEALQITREKIPKRYLFQPHNLNSNL 60  
 Db 110 TLEMGEFQIDGRSARGGOQSDANDAYAKALNAGSKSEALNVLRELAPKDYVLOPHNLNSNL 169  
 OY 61 DRIFDKTPEP 70  
 Db 170 DRIFDKTPEP 178

RESULT 7  
 S31875

AL1 protein - pepper rizado amarillo virus  
 C:Species: pepper rizado amarillo virus



Query Match	57.7%;	Score 209;	DB 1;	Length 358;
Best Local Similarity	52.9%;	Pred. No. 3.3e-15;		
Matches	37;	Conservative 17;	Mismatches 16;	Indels 0;
				Gaps 0;

QY 1 TLVWGEFQVDAAGCAGCOTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSNL 60  
 Db 107 TLEWGEFQIDGRSARGGQOSANDYAKALNAGSVQSLAVIREBQPKDFVLQNNHNRSL 166  
 QY 61 DRIFDKTPEP 70  
 Db 167 ERIFAKAPPP 176

## RESULT 12

gene C1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
 C:Accession: S39235  
 R:Crespi, S.; Norris, E.; Valira, A.; Bosco, D.; Accotto, G.  
 A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.  
 A:Reference number: S39233  
 A:Accession: S39235  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <CRE>  
 C:Cross-references: EMBL:Z28390; NID:q1041671; PID:q1334964  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.7%; Score 209; DB 2; Length 359;  
 Best Local Similarity 53.6%; Pred. No. 3.3e-15;  
 Matches 37; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDAAGCAGCOTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSNL 61  
 Db 111 LEWGTQIDGRSARGGQOTANDAYAKAINARSKSEALDVKOLAPRDYVLFHFNHNSMLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 KVFQVPPAP 179

## RESULT 13

AL1 protein - beet curly top virus  
 C:Species: beet curly top virus  
 C>Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to  
 A:Reference number: S28360  
 A:Accession: S28360  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 C:Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.7%; Score 209; DB 2; Length 385;  
 Best Local Similarity 52.9%; Pred. No. 3.5e-15;  
 Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDAAGCAGCOTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSNL 60  
 Db 137 TLEWGEFQIDGRSARGGQOTANDYAKALNATSIDQALQILKEQPKDFLQHNHNLNNA 196  
 QY 61 DRIFDKTPEP 70  
 Db 197 QKIFQRPDP 206

## RESULT 14

S59885  
 replication-associated protein C1 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C>Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
 d geminiviruses.  
 A:Reference number: S58346  
 A:Accession: S59885  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-360 <HON>  
 C:Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88223.1; PID:g974211  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 55.8%; Score 202; DB 2; Length 360;  
 Best Local Similarity 57.6%; Pred. No. 1.9e-14;  
 Matches 38; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 4 WGEFQVDAAGCAGCOTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSNDRI 63  
 Db 113 FGWFQIDGRSARGGQOSANDAYAEALNAGSKAALDILREKAPKDFVLQFNHNLNANDRI 172

QY 64 FDKTPE 69  
 Db 173 FTSAPAE 178

## RESULT 15

AL1 protein - tomato yellow leaf curl virus  
 N:Alternate names: C1 protein  
 C:Species: tomato yellow leaf curl virus  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 A:Reference number: A40779; MUID:92024070  
 A:Accession: D40779  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 C:Cross-references: GB:X15656; NID:962204; PIDN:CAA33688.1; PID:962207  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 55.2%; Score 200; DB 1; Length 357;  
 Best Local Similarity 62.3%; Pred. No. 3.1e-14;  
 Matches 38; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 4 WGEFQVDAAGCAGCOTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSNDRI 63  
 Db 111 FGVSQIDGRSARGGQOSANDAYAEALNAGSKSEALNITLKEKAPKDFVLQFNHNLNSNDRI 170  
 QY 64 F 64  
 Db 171 F 171

Search completed: October 10, 2002, 01:48:42  
 Job time: 560 sec



XX Hanley-Bowdoin L., Orozco BM, Kong L;  
 PT  
 DR WPI: 2000-618851/59.  
 XX

Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT

XX Claim 53; Page 42-43; 73pp; English.  
 PS  
 XX

The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are useful to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX

SQ Sequence 70 AA:

Query Match 100.0%; Score 362; DB 21; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 4e-37;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLWGEFQVDCAAAGCGGOTSNDAAAEALNASSKEEALQITREKIPKRYLFQFHNLNSNL 60  
 DB 1 tlwgefyvdgsaaagcgqtsndaaaealnasskeeaqlitrekipekylfghnlmsl 60  
 QY 61 DRIEIKTPEP 70  
 DB 61 dritkipep 70

RESULT 2  
 AAB18677  
 ID AAB18677 standard; peptide; 70 AA.  
 AC AAB18677;

XX 22-JAN-2001 (first entry)  
 DT  
 XX

Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 DE

Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX

OS Tomato golden mosaic virus.  
 OS

XX WO200054573-A1.  
 PN  
 XX

PD 21-SEP-2000.  
 PD

PF 15-MAR-2000; 2000WO-US06759.  
 PF

PR 18-MAR-1999; 99US-0125004.  
 PR

PR 09-APR-1999; 99US-0289346.  
 PR

PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA

PI Hanley-Bowdoin L., Orozco BM, Kong L;  
 PI

DR WPI: 2000-618851/59.  
 DR

XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT

XX Disclosure; Page 18; 73pp; English.  
 PS  
 XX

The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX

SQ Sequence 70 AA:

Query Match 96.4%; Score 349; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 1.6e-35;

Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TLWGEFQVDCAAAGCGGOTSNDAAAEALNASSKEEALQITREKIPKRYLFQFHNLNSNL 60  
 DB 1 tlwgefyvdgsaaagcgqtsndaaaealnasskeeaqlitrekipekylfghnlmsl 60  
 QY 61 DRIEIKTPEP 70  
 DB 61 dritkipep 70

RESULT 3  
 AAB18687  
 ID AAB18687 standard; peptide; 356 AA.  
 AC AAB18687;

XX 22-JAN-2001 (first entry)  
 DT  
 XX

Amino acid sequence of a geminivirus replication protein of TGMV.  
 DE

Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX

OS Tomato golden mosaic virus.  
 OS

XX Key Location/Qualifiers  
 FT Misc-difference 354  
 FT

XX /note= "unspecified amino acid"  
 XX

PN WO200054573-A1.  
 PN

PD 21-SEP-2000.  
 PD

PF 15-MAR-2000; 2000WO-US06759.  
 PF

PR 18-MAR-1999; 99US-0125004.  
 PR

PR 09-APR-1999; 99US-0289346.  
 PR

PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA

PI Hanley-Bowdoin L., Orozco BM, Kong L;  
 PI

DR WPI: 2000-618851/59.  
 DR

XX

Transgenic plants with increased resistance to geminivirus infection

comprise a nucleic acid construct containing a nucleic acid sequence  
 encoding a mutant AL1 protein with a mutation in the Rd binding region  
 Disclosure: Page 47-48; 73pp; English.

The present sequence represents a geminivirus replication (Rep)  
 protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 catalyzes cleavage and ligation of single-stranded DNA, and interacts  
 with other viral and host proteins. Mutants of the AL1 protein are used  
 to produce transgenic plants. The mutation in AL1 is present in a  
 ribosome binding region, and expression of mutant AL1 protein imparts  
 increased resistance to geminivirus infection in the plant. Mutant AL1  
 proteins are useful for producing plants having increased resistance or  
 reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 tomato mottle virus, tomato yellow leaf curl virus, potato leaf curl  
 virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 beet curly top virus.

Sequence 356 AA:

Query Match 96.4%; Score 349; DB 21; Length 356;  
 Best Local Similarity 95.7%; Pred. No. 1.2e-34;  
 Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGDGAAAGCCTSDNDAAEALNASSKEEALQIREKIPERYLRFPHNLNSNL 60  
 ||||| : |||||  
 Db 1 tlwvgefqvgsaragcqlsndaaealnasskeaalqirekipekylrfghnlmsnl 169

OY 61 DRIFDKTPEP 70  
 |||||  
 Db 170 drifdktp 179

#### RESULT 4

AA18685  
 ID AAB18685 standard; peptide; 70 AA.

AC AAB18685:  
 DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus: replication protein; Rep protein; AL1; transgenic plant;

KW ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
 OS Tomato golden mosaic virus.

Key Location/Qualifiers

FT Misc-difference 10 /note= "wild type residue replaced with Ala"

FT WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;

DR WPI; 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection

comprise a nucleic acid construct containing a nucleic acid sequence  
 encoding a mutant AL1 protein with a mutation in the Rd binding region  
 Claim 53; Page 46; 73pp; English.

The present sequence represents a mutant peptide, derived from a  
 geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 protein are used to produce transgenic plants. The mutation in AL1 is  
 present in a ribosome binding region, and expression of mutant AL1  
 protein imparts increased resistance to geminivirus infection in the  
 plant. Mutant AL1 proteins are useful for producing plants having  
 increased resistance or reduced sensitivity to a geminivirus such as  
 tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 virus, cotton leaf curl virus or beet curly top virus.

Sequence 70 AA:

Query Match 94.2%; Score 341; DB 21; Length 70;  
 Best Local Similarity 94.3%; Pred. No. 1.6e-34;  
 Matches 66; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGDGAAAGCCTSDNDAAEALNASSKEEALQIREKIPERYLRFPHNLNSNL 60  
 ||||| : |||||  
 Db 1 tlwvgefqvgsaragcqlsndaaealnasskeaalqirekipekylrfghnlmsnl 60

OY 61 DRIFDKTPEP 70  
 |||||  
 Db 61 drifdktp 70

#### RESULT 5

AA18688  
 ID AAB18688 standard; peptide; 70 AA.

AC AAB18688:  
 DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus: replication protein; Rep protein; AL1; transgenic plant;

KW ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
 OS Tomato golden mosaic virus.

Key Location/Qualifiers

FT Misc-difference 19 /note= "wild type residue replaced with Ala"

FT Misc-difference 20 /note= "wild type residue replaced with Ala"

FT WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;

DR WPI; 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rd binding region  
 PS  
 XX  
 Disclosure: Page 48; 73pp; English.

CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC plant imparts increased resistance to geminivirus infection in the  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.

SO Sequence 70 AA;

Query Match 93.4%; Score 338; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 3,6e-34;  
 Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLVWGEFPOVDGAAAGCOTSDNDAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 60  
 Db 1 tlvwgefpydgsarqcaasndaaaealnasskeeaqlirekipekyllfghlnsl 60

QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 6  
 AAB18692

ID AAB18692 standard; peptide: 70 AA.

AC AAB18692;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 ribosome binding region; resistance; geminivirus infection.

OS Synthetic;  
 OS Tomato golden mosaic virus.

Key Location/Qualifiers

FT MISC-difference 66 /note= "wild type residue replaced with Ala"

FT MISC-difference 69 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rd binding region  
 PS  
 XX  
 Disclosure: Page 50; 73pp; English.

CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC plant imparts increased resistance to geminivirus infection in the  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.

SO Sequence 70 AA;

Query Match 93.1%; Score 337; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 4,8e-34;  
 Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLVWGEFPOVDGAAAGCOTSDNDAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 60  
 Db 1 tlvwgefpydgsarqcaasndaaaealnasskeeaqlirekipekyllfghlnsl 60

QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 7  
 AAB18684

ID AAB18684 standard; peptide: 70 AA.

AC AAB18684;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 ribosome binding region; resistance; geminivirus infection.

OS Synthetic;  
 OS Tomato golden mosaic virus.

Key Location/Qualifiers

FT MISC-difference 7 /note= "wild type residue replaced with Ala"

FT MISC-difference 8 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.



```
FT      /note="wild t
```

/note="wild type residue replaced with 11"



```

FT      /note= "wild type residue replaced with Ala"
XX      WO200054573-A1.
XX      21-SEP-2000.
XX      15-MAR-2000; 2000WO-US06759.
XX      18-MAR-1999; 99US-0125004.
XX      09-APR-1999; 99US-0289346.
XX      (UYNC-) UNIV NORTH CAROLINA STATE.
XX      Hanley-Bowdoin L, Orozco BM, Kong L;
XX      WPI: 2000-618851/59.
XX      Transgenic plants with increased resistance to geminivirus infection
XX      comprise a nucleic acid construct containing a nucleic acid sequence
XX      encoding a mutant A11 protein with a mutation in the Rb binding region
XX      -
XX      Claim 52: Page 43-44; 73pp: English.
XX      The present sequence represents a mutant peptide, derived from a
XX      geminivirus replication (Rep) protein, also known as A11. A11 binds
XX      double-stranded DNA, catalyzes cleavage and ligation of single-stranded
XX      DNA, and interacts with other viral and host proteins. Mutants of the A11
XX      protein are used to produce transgenic plants. The mutation in A11 is
XX      present in a ribosome binding region, and expression of mutant A11
XX      protein imparts increased resistance to geminivirus infection in the
XX      plant. Mutant A11 proteins are useful for producing plants having
XX      increased resistance or reduced sensitivity to a geminivirus such as
XX      tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX      virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX      cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX      virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX      virus, cotton leaf curl virus or beet curly top virus.
XX      Sequence 70 AA:
XX      SO
XX      Query Match          91.4%; Score 331; DB 21; Length 70;
XX      Best Local Similarity 91.4%; Pred. No. 2.7e-33;
XX      Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
XX      QY      1 TLVWGEFVDCGAAAGCOTSDNDAAEALNASSKEEALQIIRKIPKYLFOFHNSNL 60
XX      1 TLVWGEFVDCGAAAGCOTSDNDAAEALNASSKEEALQIIRKIPKYLFOFHNSNL 60
XX      Db      1 TLVWGEFVDCGAAAGCOTSDNDAAEALNASSKEEALQIIRKIPKYLFOFHNSNL 60
XX      QY      61 DRIFDKTRPEP 70
XX      61 drifdktrpep 70
XX      Db      61 drifdktrpep 70
XX      RESULT 12
XX      AAB18691
XX      ID AAB18691 standard; peptide: 70 AA.
XX      AC AAB18691;
XX      XX
XX      DT 22-JAN-2001 (first entry)
XX      DE Mutant peptide derived from amino acids 110-179 of Rep (A11) protein.
XX      XX
XX      KM Geminivirus: replication protein; Rep protein; A11; transgenic plant;
XX      ribosome binding region; resistance; geminivirus infection.
XX      XX
XX      OS Synthetic.
XX      KM tomato golden mosaic virus.
XX      XX
XX      Key Location/Qualifiers
XX      FT MISC-difference 34

```

```

FT      /note= "wild type residue replaced with Ala"
XX      MISC-difference 35
XX      FT /note= "wild type residue replaced with Ala"
XX      FT MISC-difference 36
XX      FT /note= "wild type residue replaced with Ala"
XX      WO200054573-A1.
XX      21-SEP-2000.
XX      15-MAR-2000; 2000WO-US06759.
XX      18-MAR-1999; 99US-0125004.
XX      09-APR-1999; 99US-0289346.
XX      (UYNC-) UNIV NORTH CAROLINA STATE.
XX      Hanley-Bowdoin L, Orozco BM, Kong L;
XX      WPI: 2000-618851/59.
XX      Transgenic plants with increased resistance to geminivirus infection
XX      comprise a nucleic acid construct containing a nucleic acid sequence
XX      encoding a mutant A11 protein with a mutation in the Rb binding region
XX      -
XX      Disclosure: Page 49; 73pp: English.
XX      The present sequence represents a mutant peptide, derived from a
XX      geminivirus replication (Rep) protein, also known as A11. A11 binds
XX      double-stranded DNA, catalyzes cleavage and ligation of single-stranded
XX      DNA, and interacts with other viral and host proteins. Mutants of the A11
XX      protein are used to produce transgenic plants. The mutation in A11 is
XX      present in a ribosome binding region, and expression of mutant A11
XX      protein imparts increased resistance to geminivirus infection in the
XX      plant. Mutant A11 proteins are useful for producing plants having
XX      increased resistance or reduced sensitivity to a geminivirus such as
XX      tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX      virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX      cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX      virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX      virus, cotton leaf curl virus or beet curly top virus.
XX      Sequence 70 AA:
XX      SO
XX      Query Match          91.4%; Score 331; DB 21; Length 70;
XX      Best Local Similarity 91.4%; Pred. No. 2.7e-33;
XX      Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
XX      QY      1 TLVWGEFVDCGAAAGCOTSDNDAAEALNASSKEEALQIIRKIPKYLFOFHNSNL 60
XX      1 TLVWGEFVDCGAAAGCOTSDNDAAEALNASSKEEALQIIRKIPKYLFOFHNSNL 60
XX      Db      1 TLVWGEFVDCGAAAGCOTSDNDAAEALNASSKEEALQIIRKIPKYLFOFHNSNL 60
XX      QY      61 DRIFDKTRPEP 70
XX      61 drifdktrpep 70
XX      Db      61 drifdktrpep 70
XX      RESULT 13
XX      AAB18681
XX      ID AAB18681 standard; peptide: 70 AA.
XX      AC AAB18681;
XX      XX
XX      DT 22-JAN-2001 (first entry)
XX      DE Mutant peptide derived from amino acids 110-179 of Rep (A11) protein.
XX      XX
XX      KM Geminivirus: replication protein; Rep protein; A11; transgenic plant;
XX      ribosome binding region; resistance; geminivirus infection.
XX      XX
XX      OS Synthetic.
XX      FT MISC-difference 34

```

110-1/9 of Rep (All) protein

AAB18682;

Job time: 527 sec

```

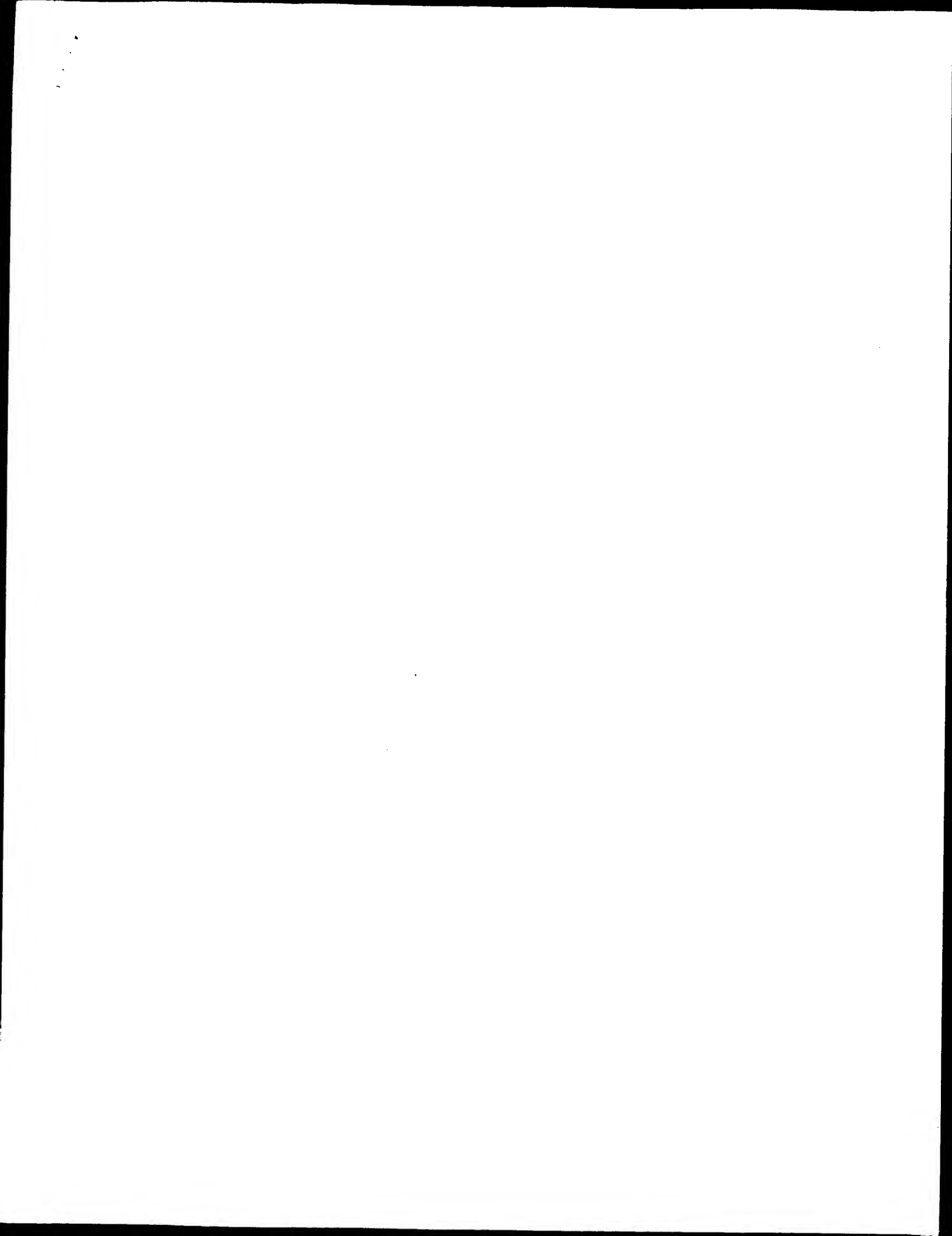
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
XX
FH Key location/Qualifiers
FT Misc-difference 52 /note= "wild type residue replaced with Ala"
FT Misc-difference 54 /note= "wild type residue replaced with Ala"
FT Misc-difference 55 /note= "wild type residue replaced with Ala"
XX
XX WO200054573-A1.
PN
XX
XX 21-SEP-2000.
PD
XX
XX 15-MAR-2000; 2000WO-US06759.
PE
XX
XX 18-MAR-1999; 99US-0125004.
PR
XX 09-APR-1999; 99US-0289346.
XX
XX (UYN-C) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
PI
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
XX Claim 53; Page 44-45; 73pp; English.
PS
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
XX
SQ Sequence 70 AA:

Query Match 89.8%; Score 325; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1.5e-32;
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVNGEFOVGAAAGCGCOTSDAAAFALNASSKEALQIRREKIPKXYLFOFHNLNSNL 60
   |||
Db 1 TLVNGEFTVQGRSARSGQLSNDAAEAALNASSKEALQIRREKIPKXYLFAFAALNSNL 60
   |||
QY 61 DRIFDKTPEP 70
   |||
Db 61 drifdktppep 70

```

Search completed: October 10, 2002, 01:47:18



Thu Oct 10 16:05:48 2002

us-09-289-346a-3.rsp

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:50 ; Search time 34.62 seconds  
(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346a-3

Perfect score: 358  
Sequence: 1 TLVWGEFYDGSARGGCT.....FOFHNLNSLDRIDPEKPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332	92.7	352	1 VAL1_TGMV	P03567 tomato gold
2	243	67.9	361	1 VAL1_PYMV	P27258 potato yell
3	230	64.2	358	1 VAL1_CLVX	P14982 cassava lat
4	230	64.2	358	1 VAL1_CLVX	P14972 cassava lat
5	222	62.0	362	1 VAL1_TYLCV	P36279 tomato yell
6	217	60.6	355	1 VAL1_ABMV	P21947 abutilon mo
7	216	60.3	349	1 VAL1_PHV	P06923 pepper huas
8	215	60.3	359	1 VAL1_TYLCV	P38609 tomato yell
9	215	60.1	361	1 VAL1_TYLCV	P05175 bean golden
10	214	59.8	357	1 VAL1_TYLCV	P06657 tomato yell
11	212	59.2	358	1 VAL1_TYLCV	P27260 tomato yell
12	208	58.1	357	1 VAL1_TYLCV	P14991 beet curly
13	190	53.1	347	1 VAL1_TYLCV	P27259 tomato yell
14	115	32.1	357	1 VAL1_TYLCV	P29048 squash leaf
15	69	19.3	577	1 T2FA_DROME	P05913 drosophila
16	66	18.4	207	1 T2FA_DROME	P35736 marmota mon
17	66	18.4	308	1 CROC_DROME	P32027 drosophila
18	65	18.2	671	1 HMOX_DROME	P22810 drosophila
19	62.5	17.5	1077	1 HLES_DROME	P02308 drosophila
20	60.5	16.9	829	1 E74B_DROME	P21005 drosophila
21	60.5	16.9	883	1 E74B_DROME	P11536 drosophila
22	60	16.8	314	1 PMXB_HUMAN	P09453 homo sapien
23	60	16.8	314	1 PMXB_HUMAN	P03569 mus musculu
24	59.5	16.6	470	1 RIBA_RHME	P02312 ribozium m
25	59	16.5	379	1 CTB_ASTE	P03818 asterina pe
26	59	16.5	498	1 CALR_PIG	P25117 sus scrofa
27	59	16.5	513	1 HEMO_CHICK	P18080 gallus gall
28	59	16.5	513	1 HEMO_CHICK	P12235 bordetella
29	58.5	16.3	1264	1 MOIR_ECOLI	P33345 escherichia
30	58.5	16.3	1713	1 LMA3_HUMAN	P16787 homo sapien
31	58	16.2	316	1 MBL_DROME	P16011 drosophila
32	58	16.2	732	1 CTG4_HUMAN	P33402 homo sapien
33	58	16.2	1081	1 U52_HSVB	P28962 equine herp

34	58	16.2	1992	1 TRIB_HUMAN	O14669 homo sapien
35	57.5	16.1	303	1 ALB3_MOUSE	P10593 zea mays (m
36	57.5	16.1	338	1 HXB_HUMAN	P31277 homo sapien
37	57.5	16.1	557	1 SASB_ANAPL	O04791 anas platyr
38	57.5	16.1	593	1 DRIL_HUMAN	O09856 homo sapien
39	57.5	16.1	601	1 DRIL_MOUSE	O62431 mus musculu
40	57.5	16.1	1556	1 PRO_DROVI	O90641 drosophila
41	57	15.9	387	1 VE2_HPV41	P27552 human papil
42	57	15.9	401	1 HME1_MOUSE	P09065 mus musculu
43	57	15.9	3080	1 POLG_ZTRVC	P18479 z genome po
44	56.5	15.8	579	1 KCN2_HUMAN	O91251 homo sapien
45	56.5	15.8	664	1 PLB1_YEAST	P39105 saccharomyc

#### ALIGNMENTS

RESULT 1	VAL1_TGMV	STANDARD:	PRT: 352 AA.
ID	VAL1_TGMV		
AC	P03567:		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	01-JUN-1994 (Rel. 29, Last annotation update)		
DE	ALI protein.		
GN	AC1		
OS	Tomato golden mosaic virus (TGMV).		
OC	Viruses; ssDNA viruses; Gemnivirus; Begomovirus.		
NCBI_TaxID=10831:			
RP	SEQUENCE FROM N.A.		
RA	Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.:		
RT	"Complete nucleotide sequence of the infectious cloned DNA components		
RT	of tomato golden mosaic virus: potential coding regions and regulatory		
RT	sequences."		
EMBO J. 3:2197-2205(1984).			
RL	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.		
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).		
DR	EMBL: K02029; .; NOT_ANNOTATED_CDS.		
DR	PIR: A04170; OOCVL.		
DR	InterPro: IPR001191; Gemin1_AL1.		
DR	Pfam: PF00799; Gemin1_AL1.		
DR	PRINTS: PR00227; GEMIN1_AL1.		
DR	ProDom: PD000736; Gemin1_AL1.		
KW	ATP-binding.		
FT	NP_BIND 223		
SEQUENCE	352 AA; 40332 MW; C33C938B5644B4A4 CRC64;		

Query Match 92.7%; Score 332; DB 1; Length 352;  
Best Local Similarity 92.9%; Pred. NO. 3.9e-31;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1 TLVWGEFYDGSARGGCTSDAAALAAALQITREKIPKYLFOFHNLNSL 60
DB	111 TLVWGEFYDGSARGGCTSDAAALAAALQITREKIPKYLFOFHNLNSL 170
QY	61 DRIFDKTPEP 70
DB	171 DRIFDKTPEP 180

RESULT 2	VAL1_PYMV	STANDARD:	PRT: 361 AA.
ID	VAL1_PYMV		

```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 protein.
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9131403; PubMed=1856690;
RA Coultas R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RT potato yellow mosaic virus."
RL J. Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL, D00940; BAA00782.1;
CC PIR: J00364; OOCVPT.
CC InterPro: IPR001191; Gemin1_AL1.
CC Pfam: PF00799; Gemin1_AL1.
CC PRINTS: PR00227; GEMCOATAL1.
CC ProDom: PD000736; Gemin1_AL1.
CC ATP-binding.
FT NP_BIND 222 229 ATP (POTENTIAL).
SO SEQUENCE 361 AA; 40850 MW; 3627A3BF1264383 CRC64;

Query Match
Best Local Similarity 67.9%; Score 243; DB 1; Length 361;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGFEFVYDGRSARGGCGTSDNDAALAAASSAAALQIIRKIPKYLFPFHNLNSNL 60
DB 110 TLWGFEFVYDGRSARGGCGTSDNDAALAAASSAAALQIIRKIPKYLFPFHNLNSNL 169
QY 61 DRIEDKTPPE 69
DB 170 DRIEDKTPPE 178

RESULT 3
VAL1_CLVN STANDARD; PRT; 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
OC AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Gay M.R.;
RT "Nucleotide sequence of cassava latent virus DNA."
RL Nature 301:260-262(1983).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL, J02057; -; NOT_ANNOTATED_CDS.
CC InterPro: IPR001191; Gemin1_AL1.
CC Pfam: PF00799; Gemin1_AL1.
CC PRINTS: PR00227; GEMCOATAL1.
CC ProDom: PD000736; Gemin1_AL1.
CC ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SO SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match
Best Local Similarity 64.2%; Score 230; DB 1; Length 358;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLWGFEFVYDGRSARGGCGTSDNDAALAAASSAAALQIIRKIPKYLFPFHNLNSNL 60
DB 109 TLWGFEFVYDGRSARGGCGTSDNDAALAAASSAAALQIIRKIPKYLFPFHNLNSNL 168
QY 61 DRIEDKTPPE 70
DB 169 DRIEDKTPPE 178

RESULT 4
VAL1_CLVN STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
OC AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain)."
RL Nucleic Acids Res. 18:197-198(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL, X17095; CA34953.1;
CC PIR: S07594; S07594.
CC InterPro: IPR001191; Gemin1_AL1.
CC Pfam: PF00799; Gemin1_AL1.
CC PRINTS: PR00227; GEMCOATAL1.
CC ProDom: PD000736; Gemin1_AL1.
CC ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SO SEQUENCE 358 AA; 40435 MW; 1DB16B80CB2D5E2C CRC64;

Query Match
Best Local Similarity 64.2%; Score 230; DB 1; Length 358;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLWGFEFVYDGRSARGGCGTSDNDAALAAASSAAALQIIRKIPKYLFPFHNLNSNL 60
DB 109 TLWGFEFVYDGRSARGGCGTSDNDAALAAASSAAALQIIRKIPKYLFPFHNLNSNL 168

```

OY 61 DRIFDKTPEP 70  
 |||| : ||  
 Db 169 DRIFPEPPAP 178

```

RESULT 5
ID VAL1_TYLCA STANDARD: PRT: 362 AA.
AC P36279:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE A1 protein (C1 protein).
CL.
OS Tomato yellow leaf curl virus (strain Australia) (TYLCAV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
CN NCBI_TaxID=36447;
RX [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93139778; PubMed=8423446;
RA Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaiian M.A.;
RT "Nucleotide sequence and genome organization of tomato leaf curl
RT geminivirus."
RL J. Gen. Virol. 74:147-151(1993)
CC -I- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.
DR PIR: J01887; J01887.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATAL1.
DR PRODOM: PD000736; Gemini_AL1.
DR ATP-binding. 221
KW NP-BIND
FT SEQUENCE 362 AA; 41197 MW; 343E7184B4704098 CRC64;
SQ

Query Match 62.0%; Score 222; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 2.2e-18;
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

OY 1 TLWGEFQVDCRSARGCCGCTSDNAAAALAASSAAALQIIRKIPKYLFOFHNLNSNL 60
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 110 TLEMGEPQIDGRSARGGQGTANDSYAKALNAGVQSALNITKEOPKDYVLONHNIRSNL 169
OY 61 DRIFDKTPEP 70
  |||| : ||
Db 170 DRIFPEPPAP 178
  |||| : ||

RESULT 6
ID VAL1_ABMVW STANDARD: PRT: 355 AA.
AC P21947:
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE A11 protein.
CL.
OS Abutilon mosaic virus (isolate West India).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
CN NCBI_TaxID=10816;
RX [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91020984; PubMed=2219703;
RA Frischmuth T., Zimmat G., Jeske H.;
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT as well as eukaryotic features."
RL Virology 178:461-468(1990).
CC -I- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: X15983; -; NOT_ANNOTATED_CDS.
DR PIR: A36214; QOCVW1.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATAL1.
DR PRODOM: PD000736; Gemini_AL1.
DR ATP-binding. 221
KW NP-BIND
FT SEQUENCE 355 AA; 40257 MW; 16A2CABA63251E95 CRC64;
SQ

Query Match 60.6%; Score 217; DB 1; Length 355;
Best Local Similarity 58.6%; Pred. No. 8.1e-18;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

OY 1 TLWGEFQVDCRSARGCCGCTSDNAAAALAASSAAALQIIRKIPKYLFOFHNLNSNL 60
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 110 TAEWGEFQIDGRSARGGQGTANDSYAKALNAGVQSALNITKEOPKDYVLONHNIRSNL 169
OY 61 DRIFDKTPEP 70
  |||| : ||
Db 170 ERIFAKAPEP 179
  |||| : ||

RESULT 7
ID VAL1_PHVU STANDARD: PRT: 349 AA.
AC O06923:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE A11 protein.
CL.
OS Pepper huasteco virus (PHV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
CN NCBI_TaxID=28349;
RX [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94015007; PubMed=8409944;
RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA Rivera-Bustamante R.F.;
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT comparison with bipartite geminiviruses."
RL J. Gen. Virol. 74:2225-2231(1993).
CC -I- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: X70418; CAA49856.1; -
DR PIR: S31875; S31875.
DR PIR: JQ2300; JQ2300.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATAL1.
DR PRODOM: PD000736; Gemini_AL1.
DR ATP-binding. 221
KW NP-BIND
FT SEQUENCE 349 AA; 39722 MW; D5F4E76CD56370F4 CRC64;
SQ

Query Match 60.3%; Score 216; DB 1; Length 349;
Best Local Similarity 58.6%; Pred. No. 1e-17;

```

Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIREKIPKYLFOFHNLNSNL 60  
 Db 110 TLWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIREKIPKYLFOFHNLNSNL 169

QY 61 DRIEDKTRPEP 70  
 Db 170 NRIFQTPPEP 179

RESULT 8  
 VAL1\_TYLCU STANDARD; PRT; 359 AA.

AC P38609;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE A11 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=37139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94256836; Pubmed=8198442;  
 RA Norris E., Hidalgo E., Accotto G., Moriones E.;  
 RT "High similarity among the tomato yellow leaf curl virus isolates  
 from the west Mediterranean basin: the nucleotide sequence of an  
 infectious clone from Spain.";  
 RL Arch. Virol. 135:165-170(1994).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z25751; CAA81026.1; -  
 DR PIR; S39211; S39211.  
 DR InterPro; IPR001191; Gemin1\_AL1.  
 DR Pfam; PF00799; Gemin1\_AL1; 1.  
 DR PRINTS; PR00227; GEMCOATAL1.  
 DR PRODOM; PD000736; Gemin1\_AL1; 1.  
 KW ATP-binding.  
 FT NP-BIND  
 SQ SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 60.1%; Score 216; DB 1; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 1.4e-17;  
 Matches 40; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 2 TLWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIREKIPKYLFOFHNLNSNL 61  
 Db 111 TLWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIREKIPKYLFOFHNLNSNL 170

QY 62 RIFDKTRPEP 70  
 Db 171 RVFQVPPAP 179

RESULT 9  
 VAL1\_BGMV STANDARD; PRT; 353 AA.

AC P05175;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE A11 protein (40.2 kDa protein).

GN ACl.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;  
 RT "Nucleotide sequence of bean golden mosaic virus and a model for gene  
 regulation in geminiviruses.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M10070; AAA46318.1; -  
 DR InterPro; IPR001191; Gemin1\_AL1.  
 DR Pfam; PF00799; Gemin1\_AL1; 1.  
 DR PRINTS; PR00227; GEMCOATAL1.  
 DR PRODOM; PD000736; Gemin1\_AL1; 1.  
 KW ATP-binding.  
 FT NP-BIND  
 SQ SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match 60.1%; Score 215; DB 1; Length 353;  
 Best Local Similarity 58.6%; Pred. No. 1.4e-17;  
 Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIREKIPKYLFOFHNLNSNL 60  
 Db 110 TLWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIREKIPKYLFOFHNLNSNL 169

QY 61 DRIEDKTRPEP 70  
 Db 170 ERIFKTRPEP 179

RESULT 10  
 VAL1\_TMOV STANDARD; PRT; 361 AA.

AC O06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A11 protein.  
 GN A11.  
 OS Tomato mottle virus (isolate Florida) (TMOV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93107858; Pubmed=1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 isolated from tomatoes in Florida.";  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L14460; AAC32414.1; -



DR PIR: J01870: J01870.  
 DR InterPro: IPR001191: Gemini\_AL1.  
 DR Pfam: PF00799: Gemini\_AL1.1.  
 DR PRINTS: PR00227: GEMCOATALL.  
 DR ProDom: PD000736: Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 222 229 ATP (BY SIMILARITY).  
 SQ SEQUENCE 361 AA: 40516 MW: 813865CEAC6950 CRC64:

Query Match 59.8%; Score 214; DB 1; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 1.8e-17;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFVQDGRSARCGCOTSDNAAAALAASSAAALQIIRKIPKYLFOFHNLNSL 60  
 DB 110 TIEWGDFQIDGRSARCGCQSDNDSYAKALNAGVOSALAVLEBPDPFVLQHNHNSL 169  
 QY 61 DRFDKTPPEP 70  
 DB 170 ERIFAKPEP 179

RESULT 11  
 VAL1\_TYLCV STANDARD; PRT: 359 AA.  
 AC P27260.  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92107660; PubMed=1840676;  
 RA Khayr-Pour A., Bendahmane M., Matzelt V., Accotto G.P., Crespi S.,  
 RA Gronenborn B.;  
 RT "Tomato yellow leaf curl virus from Sardinia is a  
 RT whitefly-transmitted monopartite geminivirus.";  
 RL Nucleic Acids Res. 19:6763-6769(1991).

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A1 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X61153: CAA43466.1: -;  
 DR PIR: S22593: S22593.  
 DR InterPro: IPR001191: Gemini\_AL1.  
 DR Pfam: PF00799: Gemini\_AL1.1.  
 DR PRINTS: PR00227: GEMCOATALL.  
 DR ProDom: PD000736: Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 359 AA: 40733 MW: 97174A07C93EFA7 CRC64:

Query Match 59.2%; Score 212; DB 1; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 3.1e-17;  
 Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LWGGEFVQDGRSARCGCOTSDNAAAALAASSAAALQIIRKIPKYLFOFHNLNSL 61  
 DB 111 LEMGTQIDGRSARCGCQOTANDAYAAKINAGSKSOLADYIKELAPDYLIHNNHNSL 170  
 QY 62 RIFDKTPPEP 70

DB 171 KYFQVPPAP 179

RESULT 12  
 VAL1\_BCTV STANDARD; PRT: 358 AA.  
 AC P14991.  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE A1 protein (40.8 kDa protein).  
 OS Beet curly top virus (BCTV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
 OX NCBI\_TaxID=10840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 RT beet curly top virus.";  
 RL EMBO J. 5:1761-1767(1986).

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A1 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X04144: -; NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191: Gemini\_AL1.  
 DR Pfam: PF00799: Gemini\_AL1.1.  
 DR PRINTS: PR00227: GEMCOATALL.  
 DR ProDom: PD000736: Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 222 229 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA: 40889 MW: 39A45FE3C0B9C33 CRC64:

Query Match 58.1%; Score 208; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 9e-17;  
 Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLWGEFVQDGRSARCGCOTSDNAAAALAASSAAALQIIRKIPKYLFOFHNLNSL 60  
 DB 110 TIEWGDFQIDGRSARCGCQOTANDSYAKALNTSIDQALQIKERQPDYLIQHNHNSL 169  
 QY 61 DRFDKTPPEP 70  
 DB 170 QKIFQRPDP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT: 357 AA.  
 AC P27259.  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RT with a single genomic component.";  
 RL Virology 185:151-161(1991).

```

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15565; CA33688.1; -
DR PIR: D40779; Q0CVCI.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemini_AL1.1.
DR ATP-binding.
KW NP_BIND
FT SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;
SQ
Query Match
Best Local Similarity 53.1%; Score 190; DB 1; Length 357;
Matches 37; Conservative 11; Mismatches 13; Indels 0; Gaps 0;
QY 4 GGEFQVDRSARSGCCTSDNDAALAAASSAAALQIIEKIPKYLFOFHNLSNLDRI 63
DB 111 GFSQIDGRSARSGCQSDNDAALAAASSAAALQIIEKIPKYLFOFHNLSNLDRI 170
QY 64 F 64
DB 171 F 171
RESULT 14
ID VAL1_SLCV STANDARD; PRT; 347 AA.
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE AL1 protein.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91082449; PubMed=1984668;
RA Lazarowitz S.G., Lazdins I.B.;
RT "Infectivity and complete nucleotide sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype."
RL Virology 180:58-69(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M38183; AAC32410.1; ALT_INIT.
DR PIR: C36785; Q0CVSI.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemini_AL1.1.
DR ATP-binding.
KW NP_BIND
FT SEQUENCE 347 AA; 39110 MW; AFDABDE122110E CRC64;
SQ

```

```

Query Match
Best Local Similarity 32.1%; Score 115; DB 1; Length 347;
Matches 24; Conservative 14; Mismatches 24; Indels 4; Gaps 1;
QY 5 GGEFQVDRSARSGCCTSDNDAALAAASSAAALQIIEKIPKYLFOFHNLSNLDRI 64
DB 116 GQYKSG-----GSKSNKQDVYNNAAVNAAGSAGBALDIITAGDPKTFIVYHNLNVERLF 171
QY 65 DKTEPP 70
DB 172 QKPEPP 177
RESULT 15
ID T2FA_DROME STANDARD; PRT; 577 AA.
AC Q05913; Q9V119;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription initiation factor IIF, alpha subunit (TFIIF-alpha)
DE (Transcription factor 5, large chain) (TF5a).
DE TFIIF-ALPHA OR CG10281.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93219133; PubMed=8464745;
RA Gong D.-W., Horikoshi M., Nakatani Y.;
RT "Analysis of cDNA encoding Drosophila transcription initiation factor
RT TFIIF alpha (RAP74).";
RL Nucleic Acids Res. 21:1492-1492(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=93219101; PubMed=8464716;
RA Kephart D., Price M.P., Burton Z.F., Finkelstein A.,
RA Greenblatt J., Price D.H.;
RT "Cloning of a Drosophila cDNA with sequence similarity to human
RT transcription factor RAP74."
RL Nucleic Acids Res. 21:1319-1319(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Franckoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butcher K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

```

RA Nelson D.R., Nelson X.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzo M., Plitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Sme B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter G.E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gidbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [4]  
 RP FUNCTION.  
 RX MEDLINE-94230464; Pubmed-8175788:  
 RA Kephart D., Wang B.Q., Burton Z.F., Price D.H.;  
 RT "Functional analysis of *Drosophila* factor 5 (Tf1f), a general  
 RT transcription factor.";  
 RL J. Biol. Chem. 269:13536-13543(1994).  
 CC -1- FUNCTION: Tf1f is a GENERAL TRANSCRIPTION INITIATION FACTOR THAT  
 CC BINDS TO RNA POLYMERASE II AND HELPS TO RECRUIT IT TO THE  
 CC INITIATION COMPLEX IN COLLABORATION WITH Tf1b. IT PROMOTES  
 CC TRANSCRIPTION ELONGATION.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: TO OTHER TRANSCRIPTION FACTOR IIF, ALPHA SUBUNIT.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L10331; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: X66982; CAA47501.1; -;  
 DR EMBL: AE003675; AAF54125.1; -;  
 DR PIR: S30237; S30237.  
 DR TRANSFAC: T02173; -;  
 DR FlyBase: FBgn0010282; Tf1f-alpha.  
 KW Transcription regulation; DNA-binding; Nuclear protein.  
 FT CONFLICT 20 A -> R (IN REF. 1).  
 FT CONFLICT 76 E -> G (1).  
 FT CONFLICT 76 E -> A (IN REF. 1).  
 FT CONFLICT 528 P -> A (IN REF. 1).  
 SO SEQUENCE 577 AA; 64466 MW; 1E3D8E8E3C833761 CMC64;

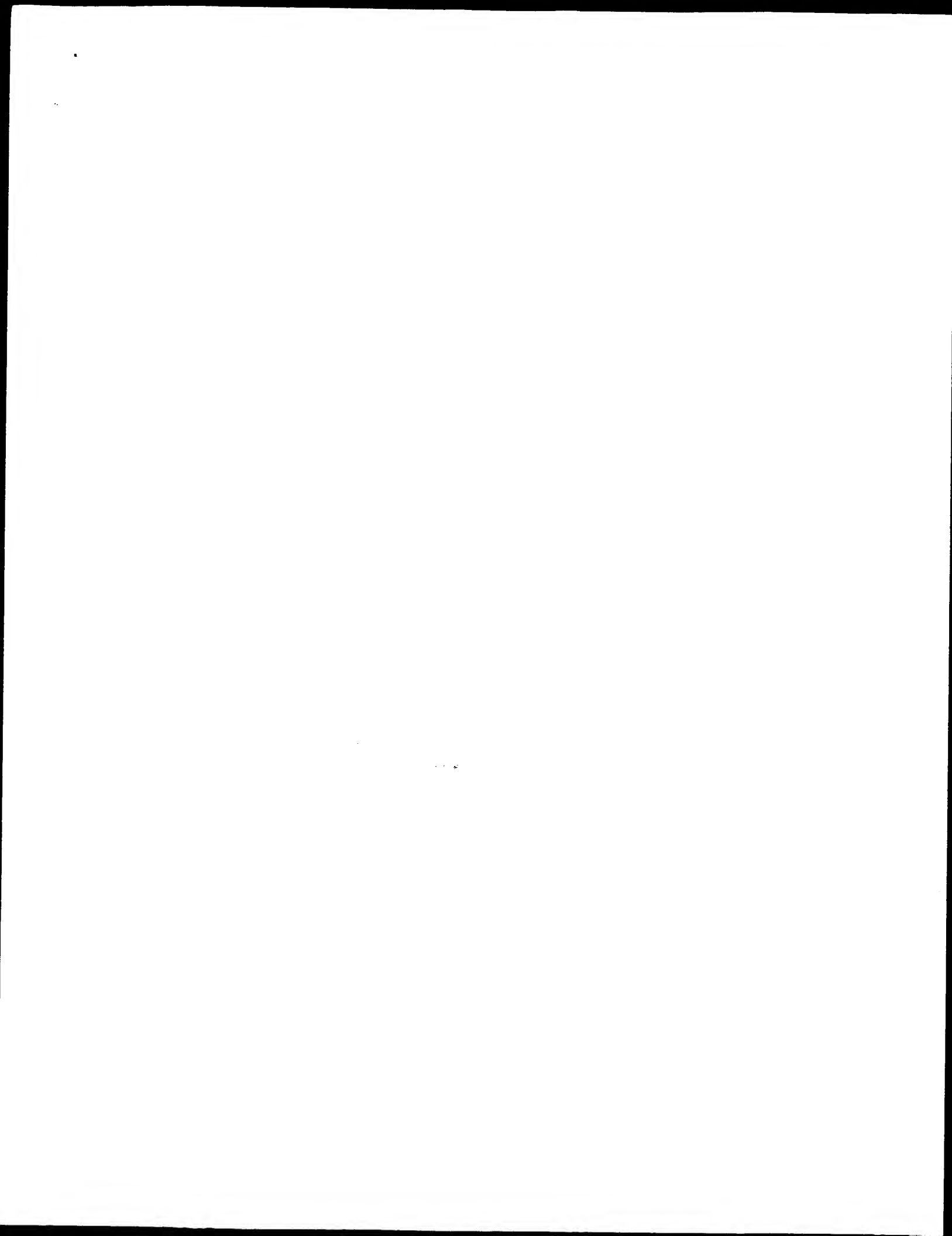
Query Match	19.38;	Score 69;	DB 1;	Length 577;
Best Local Similarity	36.48;	Pred. No. 1.9;		
Matches	20;	Conservative	9;	Mismatches 20;
			Indels	6;
			Gaps	1.

```

QY      13 SARGGCOTSDAAAAAAAAAASSAAAAAQQIIRE-----KIEKYLFQFHNLNSLND 61
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      10 SAASGSGTSAAAAAAASVASGSASSSANVOEKKIRVPKMKKKHHVMRENATLNVND 64

```

Search completed: October 10, 2002, 02:09:51  
Job time: 1349 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:42 ; Search time 69.63 Seconds  
(without alignments)  
96,600 Million cell updates/sec

Title: US-09-289-346a-3

Perfect score: 358

Sequence: 1 TLVWGEFQYDGRSARGGCT.....FOFNNLSNLDRIKTPPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR 71:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	332	92.7	352	1 QOCVLI	AL1 protein - toma
2	243	67.9	361	1 QOCVPT	AL1 protein - toma
3	230	64.2	358	2 S07594	hypothetical prote
4	222	62.0	362	1 J01887	AL1 protein - toma
5	217	60.6	355	1 QOCVWL	AV1 protein - abut
6	216	60.3	349	2 J02300	replicase - pepp
7	216	60.3	349	2 S1875	AL1 protein - pepp
8	216	60.3	359	2 S39211	gene C1 protein -
9	214	59.8	358	1 J01870	AL1 protein - toma
10	213	59.5	351	2 J02327	AL1 protein - indi
11	212	59.2	359	2 S22593	hypothetical prote
12	208	58.1	385	2 S28360	AL1 protein - beet
13	207	57.8	360	2 S59885	replication-associ
14	203	56.7	359	2 S39235	gene C1 protein -
15	190	53.1	357	1 QOCVCI	AL1 protein - toma
16	185	51.5	331	1 S45059	AC1 protein (clone
17	115	32.1	347	1 QOCVSI	AL1 protein - squa
18	69	19.3	577	2 S30237	transcription init
19	66	18.4	508	2 S59870	fork head domain P
20	66	18.4	840	2 T36175	probable large ATP
21	65	18.2	254	2 F97486	reco-like protein
22	65	18.2	254	2 AF2704	DNA repair protein
23	63	18.2	671	2 A35912	homeotic protein o
24	63	17.6	392	2 T44362	acetyl-CoA acetyl
25	62.5	17.5	700	2 T05841	spliceosome-associ
26	62.5	17.5	1077	2 A44067	serine-rich protei
27	62	17.3	409	2 T16781	hypothetical prote
28	61.5	17.2	317	2 D89661	conserved hypothet
29	61.5	17.2	327	2 F83408	hypothetical prote

30	61.5	17.2	705	2 T31157	hypothetical prote
31	61.5	17.2	1541	2 T02831	AAA protein L4171.
32	61	17.0	339	2 T26328	hypothetical prote
33	61	17.0	371	2 B69502	alcohol dehydrogen
34	61	17.0	1265	2 T47626	structural mainlen
35	61	17.0	1510	2 T33100	hypothetical prote
36	60.5	16.9	509	2 T02942	O-succinylthiomoseri
37	60.5	16.9	829	2 A34692	ecdysone-induced p
38	60.5	16.9	883	2 S04722	puff 74E protein -
39	60	16.8	314	2 JC5273	paired type homeob
40	59.5	16.6	201	2 A81380	probable flagellar
41	59.5	16.6	244	2 B86197	hypothetical prote
42	59.5	16.6	470	2 T46814	gamma-aminobutyral
43	59.5	16.6	470	2 B95419	diaminobutyrate--p
44	59.5	16.6	481	2 A70091	probable phosphoes
45	59.5	16.6	509	2 T02940	O-succinylthiomoseri

#### ALIGNMENTS

RESULT 1

QOCVLI 1

AL1 protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1352 <HAM>

A:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 92.7%; Score 332; DB 1; Length 352;

Best local similarity 92.9%; Pred. No. 4.8e-31;

Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQYDGRSARGGCTSDNDAALAAASAALQIIRKIPKYLFOFNNLSNL 60

Db 111 TLVWGEFQYDGRSARGGCTSDNDAALAAASAALQIIRKIPKYLFOFNNLSNL 170

QY 61 DRIFDKTPPEP 70

Db 171 DRIFDKTPPEP 180

RESULT 2

QOCVPT 2

AL1 protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C:Accession: J00364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: J00362; MUID:91311403

A:Accession: J00364

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-361 <COU>

A:Cross-references: GB:D00940; MID:9222458; PIDN:BAA00782.1; PID:9222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein



A;Molecule type: DNA

Query Match	59.28;	Score 212;	DB 2;	Length 359;
Best Local Similarity	56.58;	Pred. No. 5.2e-17;		
Matches 39; Conservative	10;	Mismatches 20;	Indels 0;	Gaps 0;

Oy	2	IWMGEFYVDGSSAGCGCOTSNDAAAALAASSAAAAAIOIIRKIPERKYTFQEFHINLSND	61
Dd	111	LEMGTGFIDGRSARSGGGQTANDAVAKAINMGSRSQALDVIKELAPROYVYLFHNINSND	170
Oy	62	RIFEDKTEP	70
		: : :	
Dd	171	KVFQVPFAP	179

RESULT 12

AL1 protein - beet curly top virus  
 C:Species: beet curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Stanley, J., Matham, P.G.: Callis, R.J.: Pinner, M.S.  
 EMBL J. 5, 1761-1767, 1986  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus  
 A:Reference number: S28360  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GB:M24597, EMBL:X04144, NID:g210678, PIDN:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	58.1%;	Score 208;	DB 2;	Length 385;
Best Local Similarity	55.7%;	Pred. No. 1.6e-16;		
Matches 39; Conservative	13;	Mismatches 18;	Indels 0;	Gaps 0;

```

QY      1  TLWEEFPOVDGSSANGGCOTSDNDAALAAALSSAAALAIITREKIPKEYLTLEQFINLSNL 60
Db      137  TLWEEFPOVDGSSANGGOTANDSYAKALNATSLDQALQILKEFQPKYFIQHNNLLNNA 196

QY      61  DRIFDKTREP 70
Db      197  QKIFQRPPDP 206

```

RESULT 13  
CE0000

replication-associated protein C1 - tomato yellow leaf curl virus  
 C.Species: tomato yellow leaf curl virus  
 C.Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C.Accession: S59885  
 R.Hong, Y.; Harrison, B.D.  
 submitted to the EMBL Data Library, February 1995  
 A.Description: Nucleotide sequences from tomato leaf curl viruses from different countries  
 d geminiviruses  
 A.Reference number: S58346  
 A.Accession: S59885  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-360 <HON>  
 A.Cross-references: EMBL:Z48182; NTD:g944838; PTDN:CAA88229.1; PTD:g974211  
 C.Superfamily: tomato golden mosaic virus AL1 protein

Query Match	57.8%	Score 207	DB 2	Length 360
Best Local Similarity	60.6%	Pred	No. 2e-16	
Matches 40	Conservative 12	Mismatches 14	Indels 0	Gaps 0

[illegible]

RESULT 14

gene C1 protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
C:Accession: S39235  
R:Crespi, S.; Norris, E.; Valira, A.; Bosco, D.; Accotto, G.  
submitted to the EMBL Data library, December 1993  
A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity  
A:Reference number: S39233  
A:Accession: S39235  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-359 <CRE>  
A:Cross-References: EMBL:Z28390; NID:q1041671; PID:q1334964  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	56.7%;	Score 203;	DB 2;	Length 359;
Best Local Similarity	53.6%;	Pred. No. 5.8e-16;		
Matches 37; Conservative	12;	Mismatches 20;	Indels 0;	Gaps 0;

[illegible]

RESULT	15
DOCVCT	

AL1 protein -tomato yellow leaf curl virus  
 N:Alternate names: CL protein  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 30-Jun-1992 #sequence-revision 30-Jun-1992 #text-change 23-Jul-1999  
 C:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A>Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single  
 A:Reference number: M0779; MUID:92024070  
 A:Accession: D40779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 A:Cross-references: GB:XI5656; NID:g62204; PTDN:CA33668.1; PID:g62207  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	53.1%	Score 190;	DB 1;	Length 357;
Best Local Similarity	60.7%;	Pred. No. 1.9e-14;		
Matches 37; Conservative	11;	Mismatches 13;	Indels 0;	Gaps 0;

```

QY      4 WEFPVDRSRAGCGCOTSNDAALAAASAAALQIIRKIPERYLFOFHNLSNDRI 63
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      111 FVQSDIDGRSAGCGQSANDAVAEALNCGSRSEALNTERAPDYLIQFHNLSNDRI 171
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      64 F 64
        |
Db      171 F 171

```

Search completed: October 10, 2002, 01:48:43  
Job time: 561 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds  
(without alignments)  
58.394 Million cell updates/sec

Title: US-09-289-346a-3

Perfect score: 358  
Sequence: 1 TLVWGFQVVDGHSARGGCOT.....FOFINLNSLDRIPTKPEP 70

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_032802.\*  
1: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	358	100.0	70	AA18679	Mutant peptide der
2	347	96.9	70	AA18691	Mutant peptide der
3	343	95.8	70	AA18690	Mutant peptide der
4	332	92.7	70	AA18677	Peptide fragment f
5	332	92.7	356	AA18687	Amino acid sequenc
6	324	90.5	70	AA18685	Mutant peptide der
7	321	89.7	70	AA18688	Mutant peptide der
8	320	89.4	70	AA18692	Mutant peptide der
9	318	88.8	70	AA18684	Mutant peptide der
10	317	88.5	70	AA18678	Mutant peptide der
11	317	88.5	70	AA18686	Mutant peptide der

#### ALIGNMENTS

Result	ID	Score	Query Match	Length	Description
1	AA18679	358	100.0	70	Mutant peptide der
2	AA18680	347	96.9	70	Mutant peptide der
3	AA18681	343	95.8	70	Mutant peptide der
4	AA18683	332	92.7	70	Mutant peptide der
5	AA18685	332	92.7	356	Amino acid sequenc
6	AA18687	324	90.5	70	Mutant peptide der
7	AA18688	321	89.7	70	Mutant peptide der
8	AA18692	320	89.4	70	Mutant peptide der
9	AA18684	318	88.8	70	Mutant peptide der
10	AA18678	317	88.5	70	Mutant peptide der
11	AA18686	317	88.5	70	Mutant peptide der

RESULT 1  
AA18679 standard; peptide: 70 AA.  
AA18679;  
22-JAN-2001 (first entry)  
Mutant peptide derived from amino acids 110-179 of Rep (A11) protein.  
Geminivirus; replication protein; Rep protein; A11; transgenic plant;  
ribosome binding region; resistance; geminivirus infection.  
Synthetic.  
Tomato golden mosaic virus.  
Key location/Qualifiers  
Misc-difference 27 /note= "wild type residue replaced with Ala"  
Misc-difference 30 /note= "wild type residue replaced with Ala"  
Misc-difference 34 /note= "wild type residue replaced with Ala"  
Misc-difference 35 /note= "wild type residue replaced with Ala"  
Misc-difference 36 /note= "wild type residue replaced with Ala"  
WC200054573-A1.  
21-SEP-2000.  
15-MAR-2000: 2000MO-US06759.

[illegible]

PN MO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000MO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure; Page 49; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein impacts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:  
 XX  
 Query Match 95.8%; Score 343; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 3.8e-37;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDSRAGCGCOTSDAATAAALAASSAAALQIIRKIPKRYLFQPHNLSNL 60  
 DB 1 TLVWGEFQVDSRAGCGCOTSDAATAAALAASSAAALQIIRKIPKRYLFQPHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 drifdktp 70  
 DB  
 RESULT 4  
 AAB18677  
 ID AAB18677 standard; peptide: 70 AA.  
 XX  
 AC AAB18677;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 PN MO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 PD  
 PF 15-MAR-2000; 2000MO-US06759.  
 PF

XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure; Page 18; 73pp; English.  
 XX  
 CC The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein impacts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX  
 SQ Sequence 70 AA:  
 XX  
 Query Match 92.7%; Score 332; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 1e-35; Indels 0; Gaps 0;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDSRAGCGCOTSDAATAAALAASSAAALQIIRKIPKRYLFQPHNLSNL 60  
 DB 1 TLVWGEFQVDSRAGCGCOTSDAATAAALAASSAAALQIIRKIPKRYLFQPHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 drifdktp 70  
 DB  
 RESULT 5  
 AAB18687  
 ID AAB18687 standard; peptide: 356 AA.  
 XX  
 AC AAB18687;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Amino acid sequence of a geminivirus replication protein of TGMV.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 PN Key Location/Qualifiers  
 FT MISC-difference 354  
 FT /note="unspecified amino acid"  
 FT  
 XX  
 PN MO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 PD  
 PF 15-MAR-2000; 2000MO-US06759.  
 PF

PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.

XX (UNNC-) UNIV NORTH CAROLINA STATE.  
XX

PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX

DR WPI: 2000-618851/59.  
XX

XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX

PS Disclosure: Page 47-48; 73pp; English.  
XX

CC The present sequence represents a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC tomato mottle virus, tomato yellow leaf curl virus, bean golden mosaic virus,  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
XX

SO Sequence 356 AA;

Query Match 92.7%; Score 332; DB 21; Length 356;  
Best Local Similarity 92.9%; Pred. No. 8e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 TLVWGEFQVGRSARAGCGCOTSDNDAAMAAALAAASSAAALQITREKTEPKYLFQFHNLSNL 60  
DB 110 TLVWGEFQVGRSARAGCGCOTSDNDAAMAAALAAASSAAALQITREKTEPKYLFQFHNLSNL 169

OY 61 DRIEDKTPPEP 70  
DB 170 drifdktppep 179

# RESULT 6

ID AAB18685 standard; peptide; 70 AA.

AC AAB18685;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX

OS Synthetic.  
XX

OS Tomato golden mosaic virus.  
XX

PH Key Location/Qualifiers  
FT Misc-difference 10

FT /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

FT 21-SEP-2000.

PD 15-MAR-2000; 2000WO-US06759.  
XX

PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.

XX (UNNC-) UNIV NORTH CAROLINA STATE.  
XX

PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX

DR WPI: 2000-618851/59.  
XX

XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX

PS Claim 53; Page 46; 73pp; English.  
XX

CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX

SO Sequence 70 AA;

Query Match 90.5%; Score 324; DB 21; Length 70;  
Best Local Similarity 91.4%; Pred. No. 1.1e-34;  
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 TLVWGEFQVGRSARAGCGCOTSDNDAAMAAALAAASSAAALQITREKTEPKYLFQFHNLSNL 60  
DB 1 TLVWGEFQVGRSARAGCGCOTSDNDAAMAAALAAASSAAALQITREKTEPKYLFQFHNLSNL 60

OY 61 DRIEDKTPPEP 70  
DB 61 drifdktppep 70

# RESULT 7

ID AAB18688 standard; peptide; 70 AA.

AC AAB18688;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX

OS Synthetic.  
XX

OS Tomato golden mosaic virus.  
XX

PH Key Location/Qualifiers  
FT Misc-difference 19

FT /note= "wild type residue replaced with Ala"

FT Misc-difference 20

FT /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

FT 21-SEP-2000.  
XX

PF 15-MAR-2000; 2000WMO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX  
PA (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX  
PI Hanley-Bowdoin L, Orozco BM, Kong L;  
DR WPI: 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT -  
PS Disclosure; Page 48; 73pp; English.

The present sequence represents mutant peptide, derived from a geminivirus replication (rep) protein, also known as AIL1. AIL1 binds double-stranded DNA, catalyses cleavage and ligation of single-stranded DNA, and interacts with other viral and host proteins. Mutants of the AIL1 protein are used to produce transgenic plants. The mutation in AIL1 is present in a ribosome binding region, and expression of mutant AIL1 protein imparts increased resistance to geminivirus infection in the plant. Mutant AIL1 proteins are useful for producing plants having increased resistance or reduced sensitivity to geminiviruses such as tomatillo golden mosaic virus, tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl virus, African cassava mosaic virus, Indian cassava mosaic virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or beet curly top virus.

Sequence 70 AA;

Query Match	89.7%	Score 321	DB 21	Length 70
Similarity		Pred. No. 2.8e-34		
Best Local	90.0%			
Matches 63	Conservative	0	Mismatches 7	Indels 0
				Gaps 0

**Qy** 1 TLVWGEVDYNGRARGCGCOTSNRAAAALTLAASSAAALOIREKPIPEXILPFQHNLNSLT. 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
**Db** 1 TLVWGELGVGYSIRAGYCAASNDAADEALINASKCEALDIITREKIPKYLFGFHLNLSL 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | |

## RESULT 8

ID AAB18692 standard; peptide; 70 AA.

AC AAB18692;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (All) protein  
XX  
KM Geminivirus; replication protein; Rep protein; All; transgenic plant  
KW ribosome binding region; resistance; geminivirus infection.

OS	Synthetic.
OS	Tomato golden mosaic virus.

FH	Key	Location/Qualifiers

FT	Misc-difference	66	FT	note="wild type residue replaced with Ala"

	/note= "wild type residue replaced with Ala"
FT	

PN WO2000054573-A1.

PD 21-SEP-2000.  
 XX  
 PE 15-MAR-2000: 2000MO-US06759.  
 XX  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI: 2000-618851/59.  
 XX  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PR comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant A11 protein with a mutation in the Rb binding region  
 PT -  
 XX  
 XX  
 PS Disclosure: Page 50; 73pp. English.

CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.

Sequence	70 AA
SQ	

Query Match	89.4%	Score 320:	DB 21:	length 70:
Best Local Similarity	90.0%	Pred. No. 3.8e-34:		
Matches 63: Conservative	0:	Mismatches 7:	Indels 0:	Gaps 0:

Oy	1	TlVwEeFQVDGRSARGCCTSNAAAAAALAASSAAAALQIIRKIPKPEYLFQFINLSNLT	60
Oy			
Db	1	Ltlvweigtvdgrtsaragcqlsndaaeealnasskeeaqilirekipekyllfgfnlnsnlt	60
Oy			
Db			
Oy	61	DRIFDKTPEP 70	
Oy			
Db	61	drlfdadpap 70	
Oy			
Db			

## RESULT 9

ID	AA	18684	standard; peptide; 70	AA

AC AAB18684;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (All) protein.  
 XX Geminivirus; replication protein; Rep protein; All; transgenic plant.  
 KW ribosome binding region; resistance; geminivirus infection.

05	Synthetic.
05	Tomato golden mosaic virus

FH	Key	Location/Qualifiers

```

FT      Misc-difference 7
FT      /note="wild type residue replaced with Ala"
FT

```

FT	/note= "wild type residue replaced with Ala"
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

```

PN WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX
XX 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
XX Claim 52; Page 45; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyzes cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 88.8%; Score 318; DB 21; Length 70;
Best Local Similarity 90.0%; Pred. No. 6; 9e-34;
Matches 63; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TLVWGEFQVNGRSARGCGQTSNDAAALAAASSMAAALQIIRKIPKRYLFQPHNLSNL 60
Db 1 tlwvgeavdgrrsargcgqtsndaaalaaasskcaalqilirekipekylfghnl 60
QY 61 DRIFDKTPEP 70
Db 61 drifdkipep 70
RESULT 10
AAB18678
ID AAB18678 standard; peptide; 70 AA.
XX
XX AAB18678;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 12
XX
XX /note= "wild type residue replaced with Ala"
XX
XX Misc-difference 13

```

```

FT /note= "wild type residue replaced with Ala"
FT Misc-difference 15
FT /note= "wild type residue replaced with Ala"
XX
XX WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX
XX 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
XX Claim 53; Page 42-43; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyzes cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 88.5%; Score 317; DB 21; Length 70;
Best Local Similarity 88.6%; Pred. No. 9; 4e-34;
Matches 62; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 TLVWGEFQVNGRSARGCGQTSNDAAALAAASSMAAALQIIRKIPKRYLFQPHNLSNL 60
Db 1 tlwvgeifvqdaaaagcgqtsndaaalaaasskcaalqilirekipekylfghnl 60
QY 61 DRIFDKTPEP 70
Db 61 drifdkipep 70
RESULT 11
AAB18686
ID AAB18686 standard; peptide; 70 AA.
XX
XX AAB18686;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX Tomato golden mosaic virus.
XX

```

Key	Location/Qualifiers
FT	Misc-difference 24 /note= "wild type residue replaced with Leu"
FT	Misc-difference 25 /note= "wild type residue replaced with Leu"
FT	Misc-difference 26 /note= "wild type residue replaced with Leu"
PN	WO200054573-A1.
PD	21-SEP-2000.
XX	15-MAR-2000; 2000WO-US06759.
XX	18-MAR-1999; 99US-0125004.
PR	09-APR-1999; 99US-0289346.
XX	(UYNC-) UNIV NORTH CAROLINA STATE.
PA	Hanley-Bowdoin L, Orozco BM, Kong L;
PI	WPI: 2000-618851/59.
XX	Transgenic plants with increased resistance to geminivirus infection
PT	comprise a nucleic acid construct containing a nucleic acid sequence
PT	encoding a mutant AL1 protein with a mutation in the Rb binding region
PT	-
PS	Claim 53; Page 46; 73pp; English.
XX	The present sequence represents a mutant peptide, derived from a
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC	protein are used to produce transgenic plants. The mutation in AL1 is
CC	present in a ribosome binding region, and expression of mutant AL1
CC	protein imparts increased resistance to geminivirus infection in the
CC	plant. Mutant AL1 proteins are useful for producing plants having
CC	increased resistance or reduced sensitivity to a geminivirus such as
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC	virus, cotton leaf curl virus or beet curly top virus.
XX	Sequence 70 AA:
SQ	
Query Match	88.5%; Score 317; DB 21; Length 70;
Best Local Similarity	88.6%; Pred. NO. 9,4e-34;
Matches 62: Conservative	0; Mismatches 8; Indels 0; Gaps 0.
QY	1 TLWGERVYDGRSARGCGTNSDAAALAAALAAALQIRREKPEKYLDFQHNINSL 60
Db	1 TLVWGEIVDGRSARGCGTNSDIIALLAASSKQALQIRREKPEKYLDFQHNINSL 60
QY	61 DRIEKTREP 70
Db	61 dRIEKTREP 70
RESULT 12	
AA18689	
ID	AA18689 standard; peptide; 70 AA.
XX	
AC	AA18689;
XX	
DT	22-JAN-2001 (first entry)
XX	
DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX	Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW	ribosome binding region; resistance; geminivirus infection.

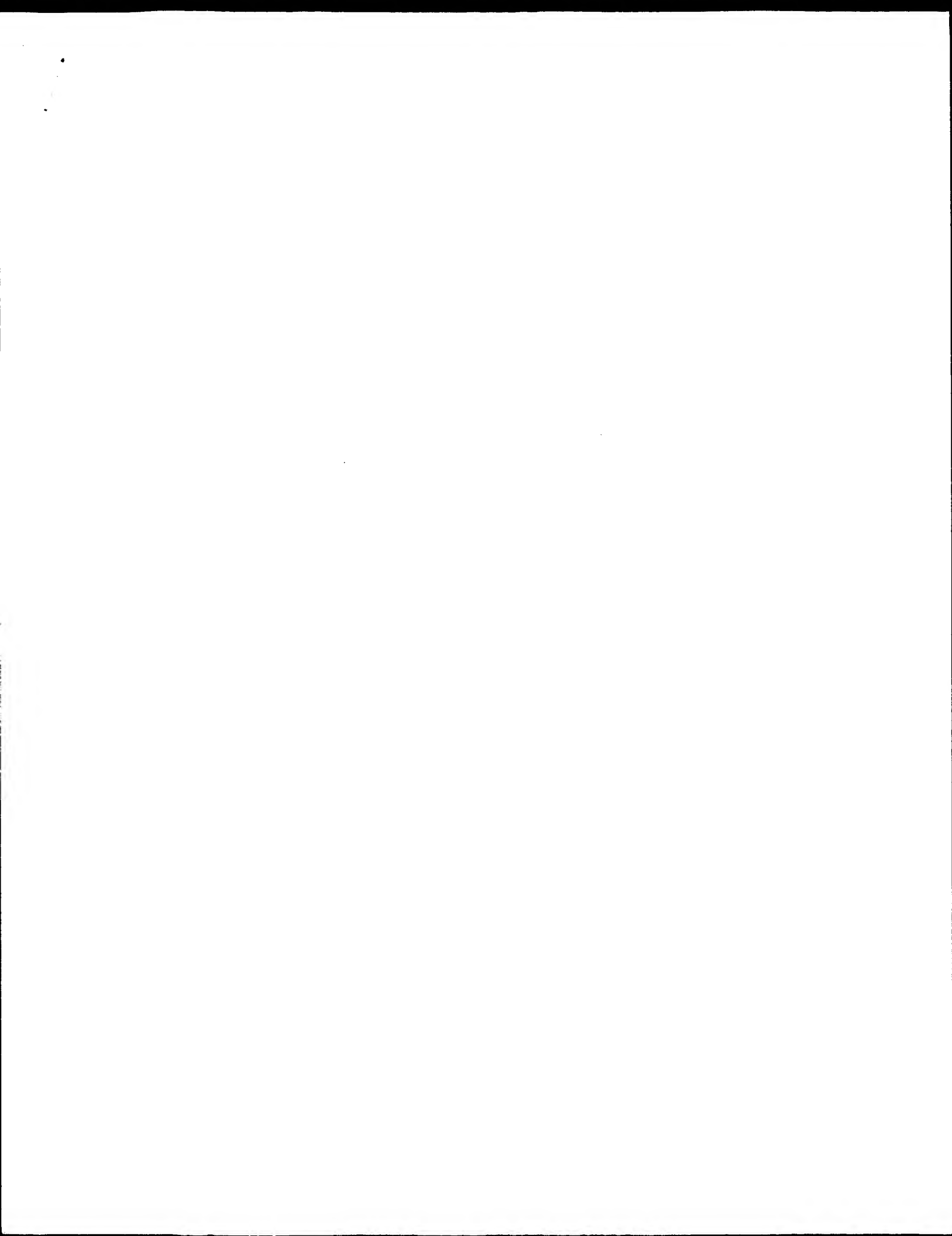
[illegible]

KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key  
 FT Misc-difference 42 Location/Qualifiers  
 FT Misc-difference 43 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 44 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 44 /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 DR WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 XX  
 PS Claim 52; Page 43-44; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:  
 Query Match 87.7%; Score 314; DB 21; Length 70;  
 Best Local Similarity 88.6%; Pred. NO. 2.3e-33;  
 Matches 62; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVGRSARSGCOTSDAAALAAASSAAALQIRKIPKXYLFQFHNLNSNL 60  
 DB 1 TLVWGEFQVGRSARSGCOTSDAAALAAASSAAALQIRKIPKXYLFQFHNLNSNL 60  
 QY 61 DRIFDKTPPEP 70  
 DB 61 DRIFDKTPPEP 70  
 RESULT 14  
 AABI8681  
 ID AABI8681 standard; peptide; 70 AA.  
 AC AABI8681;  
 XX  
 XX

DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key  
 FT Misc-difference 47 Location/Qualifiers  
 FT Misc-difference 48 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 48 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 49 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 49 /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 DR WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 XX  
 PS Claim 52; Page 44; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:  
 Query Match 86.9%; Score 311; DB 21; Length 70;  
 Best Local Similarity 88.6%; Pred. NO. 5.7e-33;  
 Matches 62; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVGRSARSGCOTSDAAALAAASSAAALQIRKIPKXYLFQFHNLNSNL 60  
 DB 1 TLVWGEFQVGRSARSGCOTSDAAALAAASSAAALQIRKIPKXYLFQFHNLNSNL 60  
 QY 61 DRIFDKTPPEP 70  
 DB 61 DRIFDKTPPEP 70  
 RESULT 15  
 AABI8683  
 AABI8683







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 seconds  
(without alignments)  
102.295 Million cell updates/sec

Title: US-09-289-346a-4  
Perfect score: 361  
Sequence: 1 TLVWGEFQVDRSARGGCGT.....FOFHNLNSLDRIFDKTPEP 70

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPTREMBL\_19.\*  
2: sp.archaea.\*  
3: sp.bacteria.\*  
4: sp.fungi.\*  
5: sp.human.\*  
6: sp.invertebrate.\*  
7: sp.mammal.\*  
8: sp.mhc.\*  
9: sp.organelle.\*  
10: sp.phage.\*  
11: sp.plant.\*  
12: sp.rodent.\*  
13: sp.virus.\*  
14: sp.vertibrate.\*  
15: sp.unclassified.\*  
16: sp.bacteriophage.\*  
17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	297	82.3	351	091R10	tomato seve
2	295	81.7	352	09E000	09e000 tomato rugo
3	281	77.8	226	029727	009727 leonurus mo
4	281	77.8	226	029727	009727 leonurus mo
5	279	77.3	361	067574	bean golden
6	277	76.7	225	090DB1	09qdb1 cowpea gold
7	274	75.9	314	09ELT8	09el18 sweet potat
8	274	75.9	364	029055	09q555 sweet potat
9	271	75.1	185	029863	09q863 sida golden
10	266	73.7	149	029895	09q895 sida golden
11	266	73.7	233	029944	09y144 macropitiliu
12	260	72.0	234	039180	039180 geminivirid
13	254	70.4	190	092089	092089 tobacco lea
14	254	70.4	190	092084	092084 tobacco lea
15	253	70.1	190	09W827	09w827 tobacco lea
16	253	70.1	208	0920C4	0920C4 tobacco lea

17	253	70.1	359	12	091M88	091m88 tobacco lea
18	251	69.5	208	12	0920C0	0920C0 tobacco lea
19	251	69.5	208	12	0920B8	0920B8 tobacco lea
20	251	69.5	363	12	072705	072705 cotton leaf
21	251	69.5	363	12	072719	072719 cotton leaf
22	249	69.0	208	12	0920B6	0920B6 tobacco lea
23	248	68.7	203	12	092083	092083 tobacco lea
24	248	68.7	359	12	092YV4	092yv4 tomato yell
25	248	68.7	359	12	092YV2	092yv2 tomato yell
26	248	68.7	359	12	09YUX7	09yux7 tomato yell
27	248	68.7	359	12	09YU27	09y127 tomato yell
28	247	68.4	359	12	088942	088942 cotton leaf
29	247	68.4	359	12	088942	088942 cotton leaf
30	245	67.9	361	12	072723	072723 cotton leaf
31	245	67.9	361	12	072723	072723 cotton leaf
32	244	67.6	348	12	0911W5	0911w5 macropitiliu
33	243	67.3	359	12	091B86	091b86 ageratum ye
34	243	67.3	360	12	090X10	090x10 ageratum ye
35	243	67.3	362	12	091M48	091m48 okra enatio
36	242	67.0	349	12	088888	088888 tomato pseu
37	242	67.0	362	12	056816	056816 chayote mos
38	241	66.8	208	12	0920A0	0920A0 south afric
39	241	66.8	354	12	091M42	091m42 tobacco lea
40	240	66.5	190	12	0920B6	0920B6 tobacco lea
41	240	66.5	208	12	0920C6	0920C6 cassava gem
42	240	66.5	358	12	09JBA2	09jba2 okra yellow
43	239	66.2	363	12	073494	073494 cotton leaf
44	237	65.7	306	12	091E71	091e71 cotton leaf
45	237	65.7	353	12	072692	072692 beet curly

## ALIGNMENTS

RESULT 1  
ID 091R10 PRELIMINARY: PRT: 351 AA.  
AC 091R10;  
DT 01-DEC-2001 (Tremblurel, 19, Created)  
DT 01-DEC-2001 (Tremblurel, 19, Last sequence update)  
DT 01-DEC-2001 (Tremblurel, 19, Last annotation update)  
DE REP PROTEIN.  
GN ACL.  
OS Tomato severe rugose virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=158463;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MINAS GERATIS;  
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.F.;  
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting begomovirus, Tomato severe rugose virus, in Brazil."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY029750; ARK50357.1; -  
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 82.3% Score 297; DB 12; Length 351;  
Best Local Similarity 78.6% Pred. No. 2.9e-27;  
Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDRSARGGCGTNDAAAEALNNSKEEALQITIAAIPKRYLFQFHNLNSL 60  
111 TLVWGEFQVDRSARGGCGTNDAAAEALNNSKEEALQITIAAIPKRYLFQFHNLNSL 170  
DB 171 DRIFARAPP 180  
QY 61 DRIFARAPP 70  
DB 171 DRIFARAPP 180  
RESULT 2  
Q9E000 PRELIMINARY: PRT: 352 AA.

AC 09E000;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN.  
 GN AC1.  
 OS Tomato rugose mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=134599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fernandes J.J., Fontes F.P.B., Brommonschenkel S.H., Carvalho M.G.,  
 RA Zambolim E.M., Zerbini F.M.,  
 RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
 RT (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,  
 RT Minas Gerais, Brazil.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF291705; AAG15546.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 SO SEQUENCE 352 AA; 40012 MW; 47CD5583BE2AD613 CRC64;

Query Match 81.7%; Score 295; DB 12; Length 352;  
 Best Local Similarity 78.6%; Pred. No. 5, 1e-27;  
 Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDRSARSGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 60  
 DB 111 TLVWGEFQVDRSARSGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 170  
 OY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 3  
 ID 009727 PRELIMINARY; PRT; 226 AA.  
 AC 009727;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonurus mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=58177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=LEWY-BRAZIL 1;  
 RA Faria J.C., Maxwell D.P.;  
 RT "Variability in geminivirus associated with Phaseolus vulgaris in  
 RT Brazil."  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U92532; AAB51157.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 226  
 FT 226  
 SO SEQUENCE 226 AA; 25617 MW; 73CD86E766083FC5 CRC64;

Query Match 77.8%; Score 281; DB 12; Length 226;  
 Best Local Similarity 78.6%; Pred. No. 1, 4e-25;  
 Matches 55; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDRSARSGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 60  
 DB 111 TLVWGEFQVDRSARSGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 170

OY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 4  
 ID 09WHF6 PRELIMINARY; PRT; 226 AA.  
 AC 09WHF6;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Tomato mild mottle geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
 OX NCBI\_TaxID=92943;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=HN96-HSKW;  
 RA Nakha M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
 RA Maxwell D.P.;  
 RT "Molecular characterization and DNA-based detection methods for  
 RT vegetable-infecting geminiviruses in Central America."  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF131071; AAD3471.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 226  
 FT 226  
 SO SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 77.8%; Score 281; DB 12; Length 226;  
 Best Local Similarity 74.3%; Pred. No. 1, 4e-25;  
 Matches 52; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDRSARSGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 60  
 DB 111 TLVWGEFQVDRSARSGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 170  
 OY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 5  
 ID 067574 PRELIMINARY; PRT; 361 AA.  
 AC 067574;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN ALL.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 RA Maxwell D.P., Russell D.R.;  
 RT "Cloning of the complete DNA genomes of four bean-infecting  
 RT geminiviruses and determining their infectivity by electric discharge  
 RT particle acceleration."  
 RL Phytopathology 81:980-985(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 RA Morales F.J., Maxwell D.P.;  
 RT "Differentiation of bean-infecting geminiviruses by nucleic acid

RT hybridization probes and aspects of bean golden mosaic in Brazil.";  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
 RT "Genetic diversity in geminiviruses causing bean golden mosaic  
 disease: The nucleotide sequence of the infectious cloned DNA  
 components of a Brazilian isolate of bean golden mosaic geminivirus";  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M88686; AAA46312.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.3%; Score 279; DB 12; Length 361;  
 Best Local Similarity 77.6%; Pred. No. 4.3e-25;  
 Matches 52; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Oy 4 WGEFQVDSRARGCGCTSDNDAAEALNASSKEEALQITIAAIPKYLFOFHNLNSNDRI 63  
 Db 113 WGEFQVDSRARGCGCTSDNDAAEALNASSKEEALQITIAAIPKYLFOFHNLNSNDRI 172  
 Oy 64 FDKTRPP 70  
 Db 173 FTKAPDP 179

RESULT 6  
 Q90DB1 PRELIMINARY; PRT; 225 AA.  
 ID Q90DB1  
 AC Q90DB1  
 DT 01-MAY-2000 (TREMUREL. 13, Created)  
 DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=69263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGMV-BR;  
 RA Faria J.C.;  
 RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from  
 Brazil";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF188708; AAF06318.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 225  
 FT SEQUENCE 225 AA; 25766 MW; 1089CB6BD8D15B5D CRC64;

Query Match 76.7%; Score 277; DB 12; Length 225;  
 Best Local Similarity 77.6%; Pred. No. 4.3e-25;  
 Matches 52; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Oy 4 WGEFQVDSRARGCGCTSDNDAAEALNASSKEEALQITIAAIPKYLFOFHNLNSNDRI 63  
 Db 113 WGEFQVDSRARGCGCTSDNDAAEALNASSKEEALQITIAAIPKYLFOFHNLNSNDRI 172  
 Oy 64 FDKTRPP 70  
 Db 173 FTKAPDP 179

RESULT 7  
 Q9ELT8

ID Q9ELT8 PRELIMINARY; PRT; 314 AA.  
 AC Q9ELT8  
 DT 01-MAR-2001 (TREMUREL. 16, Created)  
 DT 01-MAR-2001 (TREMUREL. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States";  
 RL Plant Dis. 82:1253-1257(1998).  
 DR EMBL: AF288227; AAG01006.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 75.9%; Score 274; DB 12; Length 314;  
 Best Local Similarity 79.4%; Pred. No. 1.4e-24;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Oy 1 TLWGEFQVDSRARGCGCTSDNDAAEALNASSKEEALQITIAAIPKYLFOFHNLNSNDRI 60  
 Db 110 TLWGEFQVDSRARGCGCTSDNDAAEALNASSKEEALQITIAAIPKYLFOFHNLNSNDRI 169  
 Oy 61 DRIFDKTP 68  
 Db 170 DRIFDKTP 177

RESULT 8  
 Q9QSS5 PRELIMINARY; PRT; 364 AA.  
 ID Q9QSS5  
 AC Q9QSS5  
 DT 01-MAY-2000 (TREMUREL. 13, Created)  
 DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States";  
 RL Plant Dis. 82:1253-1257(1998).  
 DR EMBL: AF104036; AAD47173.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;



Query Match	72.08;	Score 260;	DB 12;	Length 234;
Best Local Similarity	68.68;	Pred. No. 4.8e-23;		
Matches 48;	Conservative 12;	Mismatches 10;	Indels 0;	Gaps 0

RESULT	13	
Q9Z089		
ID	Q9Z089	PRELIMINARY;
CC	00Z089	PRT; 190 AA

DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TOBACCO LEAF CURL VIRUS CI AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL  
DE AND COMPLETE CDS (FRAGMENT).  
GN CI.

05 tobacco leaf curl virus.  
0C viruses: ssDNA viruses; Geminiviridae; Begomovirus  
0X NCBI\_TaxID=67762;  
11  
NN

NP SEQUENCE FROM N.A.  
 RC STRAIN=YOKOHAMA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminiviruses infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL; AB001315; BAA34033.1; -;  
 DR Interpro; IPR001191; Gemini\_AL1.  
 DR Pfam; PF00799; Gemini\_AL1.  
 DR PRINTS; PR00227; GEMCONTAL1.  
 DR ProDom; PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 1  
 FT NON\_TER 190 190  
 SO SEQUENCE 190 AA; 21432 MW; AAC093D1610FAD CRC64;

Query Match	70.48;	Score 254;	DB 12;	Length 190;
Best Local Similarity	58.98;	Word no 1	0-20	

Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY I TLWVGEOVDGRSARGCCOTSNDAAAEALNASSKEEALQITIAAIPKEYLEOFHNLNSNL 60  
||| |:||||| :||| ||||| ||| ||| :||| :||| :|||

```

Db      85 TLEWETQIDGSSAGGCONNDACAELNLSAAEALATIREKLPDQFIQYHNLSNL 144
Qy      61 DRI-----FDKTP 69
        ||: ||
Db      145 DRIAPPLEVFCPTASSFDQVE 169

```

RESULT 14

ID	Q9Z084	PRELIMINARY;	PRT;	190 AA
AC	Q9Z084			

DI	01-MAY-1999 (TREMBlREl. 10, Created)
DT	01-MAY-1999 (TREMBlREl. 10, Last Seq

TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YORK

GN cl.  
OS tobacco leaf virus

00 viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
0X NCBI\_TaxID=67762;

RP	SEQUENCE FROM N.A
PC	STPATN-YOKOHAMA

RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
RT "Molecular phylogeny of geminivirus infe-

U. FRAUENFELDER: LIV:24/-25(199/).  
EMBL; AB001318; BAA34039.1; -

```
DR      Pfam; PF00799; Gemini_ALL; 1.  
DR      PRINTS. PR00227. GEMCOATAT1
```

```
DR Probom; PB000/36; Gemini_All; 1.
FT NON_TER 1 1
```

SEQUENCE 190 AA; 2144 MW; AAC1C2943E3F01AD CRC64;

Query Match	70.4%;	Score 254;	DB 12;	Length 190;
Best Local Similarity	58.8%;	Pred. No. 1.9e-22;		
Matches 50; Conservative	8;	Mismatches 11;	Indels 16;	Gaps 1.

QY 1 TLVWGEFYDGRSARGCCQTSMDAAAEFLNASSSEALQIIAAAIPEKYLDFQHNLSNL 60  
|||:::|||||||::|:|||||||::|:|||||||  
Db 85 TLEWGTFOIDGRSARGCCQNDADCAEFLNASSKAEALITIREKLPKDFIIFYHNLSNL 144

```
QY      61 DR-----FDKTP 69
          |||
          ||: ||
Db      145 DRIFAPLEVFVCPFTASSFDQVPE 165
```

RESULT	15
Q9W827	
ID	Q9W827
PRELIMINARY:	
DPT:	100.22

DT 01-NOV-1999 (Tremblay, 12, Created)

01-DEC-2001 (Tremblay, 19, last annotation update)

DE COMPLETE CDS (FRAGMENT).  
GN C1.

Viruses; ssDNA viruses; Geminiviridae; Begomovirus

RN	[1]
RD	SEQUENCE FROM N A

RA Ooi K., Ohshita S., Ishii I., Yabara T.;  
RC STRAIN=GORAI;

J. Plant Res. 110:247-257 (1997).

DR EMBL; AB001303; BAA34010.1; -.  
DR InterPro; IPR001191; Gemini ALL.

PRINTS; PR00227; GEMCOATALL.





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:51 ; Search time 34.62 Seconds  
(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346a-4

Perfect score: 361  
Sequence: 1 TLWGEFQVDSRGSGCOT.....FOFHNLNSMLRIPDKTEEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	346	95.8	352	1 VAL1_TGMV	P03567 tomato gold
2	249	69.0	361	1 VAL1_PYMV	P27258 potato yell
3	240	66.5	358	1 VAL1_CIVK	P14982 cassava lat
4	240	66.5	358	1 VAL1_CIVN	P14972 cassava lat
5	233	64.5	362	1 VAL1_TYLCV	P36279 tomato yell
6	229	63.4	359	1 VAL1_TYLCU	P38609 tomato yell
7	222	61.5	349	1 VAL1_PHVU	P06923 pepper huas
8	222	61.5	359	1 VAL1_TYLCM	P27260 tomato yell
9	217	60.1	353	1 VAL1_BGMV	P05175 bean golden
10	216	59.8	355	1 VAL1_ABMVW	P21947 abutilon mo
11	213	59.0	358	1 VAL1_BCTV	P14991 beet curly
12	210	58.2	361	1 VAL1_TMOV	P06657 tomato molt
13	200	55.4	357	1 VAL1_TYLCV	P27259 tomato yell
14	128	35.5	347	1 VAL1_SLCV	P29048 squash leaf
15	61.5	17.0	299	1 Y175_HELPJ	Q94mq7 helicobacte
16	61.5	17.0	334	1 G3P_BACST	P00362 bacillus st
17	61.5	17.0	335	1 G3P_BACCO	P15115 bacillus co
18	61.5	17.0	335	1 G3P_BACCO	P15115 bacillus co
19	59	16.3	1483	1 BUB1_SCHPO	P12351 saccharomyc
20	58.5	16.2	492	1 RRPV_RABVP	P06747 rabies viru
21	58	16.1	492	1 MOR3_MOUSE	O35308 mus musculu
22	58	16.1	492	1 MOR3_RAT	O70461 rattus norv
23	58	16.1	555	1 GLPD_BACSU	P18158 bacillus su
24	58	16.1	601	1 CYS1_BUCAT	P57503 buchnera ap
25	57	15.8	223	1 DGR2_LACAC	O59484 lactobacill
26	57	15.8	316	1 SPDE_COFAR	O82147 coffea arab
27	57	15.8	379	1 CYB_ASTPE	Q33818 asterina pe
28	57	15.8	387	1 Y4PF_RHISN	P55615 rhizobium s
29	57	15.8	836	1 GCSR_HUMAN	O99062 homo sapien
30	57	15.8	1287	1 SKI2_YEAST	P35207 saccharomyc
31	56.5	15.7	863	1 AMPN_CAUCR	P37893 caulobacter
32	56	15.5	367	1 LHX4_MOUSE	P53776 mus musculu
33	56	15.5	577	1 PTLB_LACCA	P24400 lactobacill

34	56	15.5	874	1 SLAP_BACLI	P49052 bacillus li
35	55.5	15.4	299	1 Y175_HELPY	P56112 helicobacte
36	55.5	15.4	323	1 VAL1_PASVK	O00338 panicum str
37	55.5	15.4	397	1 METL_RAT	P13444 rattus norv
38	55.5	15.4	428	1 GBA1_CANAL	P28868 candida alb
39	55	15.2	355	1 CRTB_RHOSH	P54905 rhodobacter
40	55	15.2	524	1 MPPA_MOUSE	Q94c61 mus musculu
41	55	15.2	524	1 MPPA_RAT	P20069 rattus norv
42	55	15.2	525	1 MPPA_HUMAN	O10713 homo sapien
43	54.5	15.1	297	1 RRPV_RABVE	P22559 rabies viru
44	54.5	15.1	405	1 WCAC_ECOLI	P71237 escherichia
45	54.5	15.1	445	1 ACSC_MOOTH	Q07340 moorilla th

## ALIGNMENTS

## RESULT 1

VAL1\_TGMV STANDARD: PRT: 352 AA.

AC P03567; 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE AL1 protein.

OS Tomato golden mosaic virus (TGMV).

OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

OX NCBI\_Taxid=10831;

RN [1]

RP SEQUENCE FROM N.A. Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;

RA "Complete nucleotide sequence of the infectious cloned DNA components

RT of tomato golden mosaic virus: potential coding regions and regulatory

RT sequences." J. 3:2197-2205(1984).

RL EMBL J. 3:2197-2205(1984).

CC -I- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL, KO2029; -- NOT\_ANNOTATED\_CDS.

DR PIR: A04170; Q0CVL1.

DR InterPro: IPR001191; Gemini\_AL1.

DR Pfam: PF00799; Gemini\_AL1; 1.

DR PRINTS: PR00227; GEMCOTALL.

DR Prodom: PD000736; Gemini\_AL1; 1.

KW ATP-binding; 223 230 ATP (POTENTIAL).

FT NP\_BIND 352 AA; 40332 MW; C3C9389644B4A4 CRC64;

SO SEQUENCE 352 AA; 40332 MW; C3C9389644B4A4 CRC64;

Query Match 95.8%; Score 346; DB 1; Length 352;

Best Local Similarity 95.7%; Prod. No. 5.2e-34;

Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRGSGCOTSNDAAAEALNASSKEEALQITAAAIPEKYLFQFHNLNSL 60

DB 111 TLWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSL 170

QY 61 DRIFDKTEEP 70

DB 171 DRIFDKTEEP 180

QY 61 DRIFDKTEEP 70

DB 171 DRIFDKTEEP 180

QY 61 DRIFDKTEEP 70

DB 171 DRIFDKTEEP 180

QY 61 DRIFDKTEEP 70

DB 171 DRIFDKTEEP 180

QY 61 DRIFDKTEEP 70

DB 171 DRIFDKTEEP 180

QY 61 DRIFDKTEEP 70

DB 171 DRIFDKTEEP 180

```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 protein.
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311403; PubMed=1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RL J. Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D00940; BAA00782.1; -
DR PIR; J00364; QOCVPT.
DR InterPro; IPR001191; Geminl.AL1.
DR Pfam; PF00799; Geminl.AL1.1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRODOM; PD000736; Geminl.AL1.1.
DR ATP-binding.
KW NP-BIND.
FT 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match 69.0%; Score 249; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 2.2e-22;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGRSARGCGCOTSNDAAEALNASSKEALQITAAIPEKYLFOFHNSNL 60
DB 110 TIEWGFQIDGRSARGGQGVNDAAEALNASSGKEALMKIKRKEFLFYHNSNL 169
QY 61 DRIEDKTPPE 69
DB 170 DRIEMKAPF 178

RESULT 3
VAL1_CLV1 STANDARD; PRT; 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RX Stanley J., Gay M.R.;
RT "Nucleotide sequence of cassava latent virus DNA.";
RL Nature 301:260-262(1983).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; J02057; -; NOT ANNOTATED CDS.
DR InterPro; IPR001191; Geminl.AL1.
DR Pfam; PF00799; Geminl.AL1.1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRODOM; PD000736; Geminl.AL1.1.
DR ATP-binding.
KW NP-BIND.
FT 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 66.5%; Score 240; DB 1; Length 358;
Best Local Similarity 61.4%; Pred. No. 2.7e-21;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGRSARGCGCOTSNDAAEALNASSKEALQITAAIPEKYLFOFHNSNL 60
DB 109 TVEWGFQIDGRSARGGQGSANDAYAKALNASSKEALVIRELPKDFVLQFHNSNL 168
QY 61 DRIEDKTPPE 70
DB 169 DRIEQEPAP 178

RESULT 4
VAL1_CLV1 STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RX Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RL African cassava mosaic virus (Nigerian strain).";
RN Nucleic Acids Res. 18:197-198(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X17095; CAA34953.1; -
DR PIR; S07594; S07594.
DR InterPro; IPR001191; Geminl.AL1.
DR Pfam; PF00799; Geminl.AL1.1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRODOM; PD000736; Geminl.AL1.1.
DR ATP-binding.
KW NP-BIND.
FT 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40435 MW; 1DB16B8CB2D52C CRC64;

Query Match 66.5%; Score 240; DB 1; Length 358;
Best Local Similarity 61.4%; Pred. No. 2.7e-21;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGRSARGCGCOTSNDAAEALNASSKEALQITAAIPEKYLFOFHNSNL 60
DB 109 TVEWGFQIDGRSARGGQGSANDAYAKALNASSKEALVIRELPKDFVLQFHNSNL 168

```

OY 61 DRIFDKTPEP 70  
 DB 169 DRIFDEPPAP 178

## RESULT 5

VAL1\_TYLCU STANDARD: PRT: 362 AA.  
 AC P36279;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN Tomato yellow leaf curl virus (strain Australia) (TYLCV).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36447;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93139778; PubMed=8423446;  
 RA Dry I.B., Rygden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;  
 RT "Nucleotide sequence and genome organization of tomato leaf curl  
 geminivirus.";  
 RL J. Gen. Virol. 74:147-151(1993).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 DR PIR: J01887; J01887.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding; 221 228 ATP (POTENTIAL).  
 FT NP\_BIND  
 SQ SEQUENCE 362 AA: 41197 MW: 343E7184B4704098 CRC64;

Query Match 64.5%; Score 233; DB 1; Length 362;  
 Best Local Similarity 54.1%; Pred. No. 1.9e-20;  
 Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

OY 1 TLVWGEFQVDRSARGCGQTSNDAAEALNASSKEALQIIAAIPKYLQFPHNLSMD 60  
 DB 110 TLWGEFQIDGRSARGGQGSANDAYAOALNTGSKSEALVRELAPKDYLVQFHNLSMD 169  
 OY 61 DRI-----FDKTPPE 69  
 DB 170 DRIFTPLEVVYVSPFLSSFDVPE 194

## RESULT 6

VAL1\_TYLCU STANDARD: PRT: 359 AA.  
 AC P36609;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=37139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94256836; PubMed=8198442;  
 RA Norris E., Hidalgo E., Accotto G., Moriones E.;  
 RT "High similarity among the tomato yellow leaf curl virus isolates  
 from the west Mediterranean basin: the nucleotide sequence of an  
 infectious clone from Spain.";  
 RL Arch. Virol. 135:165-170(1994).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: 225751; CAA81026.1; -  
 DR PIR: S39211; S39211.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding; 221 228 ATP (POTENTIAL).  
 FT NP\_BIND  
 SQ SEQUENCE 359 AA: 41065 MW: 2D170A51EF80A3EC CRC64;

Query Match 63.4%; Score 229; DB 1; Length 359;  
 Best Local Similarity 60.9%; Pred. No. 5.5e-20;  
 Matches 42; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVWGEFQVDRSARGCGQTSNDAAEALNASSKEALQIIAAIPKYLQFPHNLSMD 61  
 DB 111 LEWGTFOIDGRSARGGQGTANDAYAKAINAGSKSEALDVKEKLAPKDYLVQFHNLSMD 170  
 OY 62 RIFDKTPEP 70  
 DB 171 RVQVPPAP 179

## RESULT 7

VAL1\_PHV STANDARD: PRT: 349 AA.  
 AC 006923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN Pepper huasteco virus (PHV).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=28349;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94015007; PubMed=8409944;  
 RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,  
 RA Rivera-Bustamante R.F.;  
 RT "Complete nucleotide sequence of pepper huasteco virus: analysis and  
 RT comparison with bipartite geminiviruses.";  
 RL J. Gen. Virol. 74:2225-2231(1993).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X70418; CAA49856.1; -  
 DR PIR: S31875; S31875.  
 DR PIR: J02300; J02300.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding; 221 228 ATP (BY SIMILARITY).  
 FT NP\_BIND  
 SQ SEQUENCE 349 AA: 39722 MW: D5FAE76CD56370F4 CRC64;

Query Match 61.5%; Score 222; DB 1; Length 349;

Best Local Similarity 60.0%; Pred. No. 3,7e-19;  
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARSGCQTSNDAAEALNASSKEALQITIAAIPKYLFOFHINSNL 60  
D 110 TLVWGEFQVDRSARSGCQTSNDAAEALNASSKEALQITIAAIPKYLFOFHINSNL 169  
QY 61 DRIPDKTPEP 70  
D 170 NRIFQTPPEP 179

## RESULT 8

VAL1\_TYLCM STANDARD; PRT; 359 AA.  
AC P27260;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 protein (C1 protein).  
CC  
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92107660; PubMed=1840676;  
RA Khey-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,  
RA Gronenborn B.;  
RT "Tomato yellow leaf curl virus from Sardinia is a  
RT whitefly-transmitted monopartite geminivirus.";  
RL Nucleic Acids Res. 19:6763-6769(1991).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: X61153; CAA43466.1; -  
DR PIR: S22593; S22593.  
DR InterPro: IPR001191; Gemin1\_AL1.  
DR Pfam: PF00799; Gemin1\_AL1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR ProDom: PD000736; Gemin1\_AL1.  
KW ATP-binding.  
FT NP\_BIND 220 227 ATP (POTENTIAL).  
SO SEQUENCE 359 AA; 40733 MW; 9717BRA07C93BFA7 CRC64;

Query Match 61.5%; Score 222; DB 1; Length 359;  
Best Local Similarity 58.0%; Pred. No. 3.8e-19;  
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVMGEFQVDRSARSGCQTSNDAAEALNASSKEALQITIAAIPKYLFOFHINSNL 61  
D 111 LVMGEFQVDRSARSGCQTSNDAAEALNASSKEALQITIAAIPKYLFOFHINSNL 170  
QY 62 RIFDKTPEP 70  
D 171 KVFQVPPAP 179

## RESULT 9

VAL1\_BGMV STANDARD; PRT; 353 AA.  
AC P05175;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE AL1 protein (40.2 kDa protein).  
GN ACL.

OS Bean golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10833;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;  
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene  
RT regulation in geminiviruses.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).

-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL: M10070; AAA46318.1; -  
DR InterPro: IPR001191; Gemin1\_AL1.  
DR Pfam: PF00799; Gemin1\_AL1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR ProDom: PD000736; Gemin1\_AL1.  
KW ATP-binding.  
FT NP\_BIND 222 229 ATP (POTENTIAL).  
SO SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match 60.1%; Score 217; DB 1; Length 353;  
Best Local Similarity 60.0%; Pred. No. 1.5e-18;  
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARSGCQTSNDAAEALNASSKEALQITIAAIPKYLFOFHINSNL 60  
D 110 TLVWGEFQVDRSARSGCQTSNDAAEALNASSKEALQITIAAIPKYLFOFHINSNL 169  
QY 61 DRIPDKTPEP 70  
D 170 ERIFFKPEP 179

## RESULT 10

VAL1\_ABMV STANDARD; PRT; 355 AA.  
AC P21947;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 protein.  
GN ACL.  
OS Abutilon mosaic virus (isolate West India).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91020984; PubMed=2219703;  
RA Frischmuth T., Zimmatt G., Jeske H.;  
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
RT as well as eukaryotic features.";  
RL Virology 178:461-468(1990).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL: X15983; NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; OOCVMI.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 DR ATP-binding. 221 228 ATP (POTENTIAL).  
 FT NP\_BIND  
 SQ SEQUENCE 355 AA; 40257 MW; 16A2CA8A63251E95 CRC64;

Query Match 59.8%; Score 216; DB 1; Length 355;  
 Best Local Similarity 58.6%; Pred. No. 2e-18;  
 Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

OY 1 TLWGEFQVDSRGSGCCTNSDAAAEALNASSKEEALQITAAIPEKYLFOFHNLSNL 60  
 DB 110 TLWGEFQIDGRSARGGQOTANDSYAKALNAGDVQSALNLKEOPKDYLVONHNIRSNL 169  
 OY 61 DRIFDKTPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 11  
 VAL1\_BCTV STANDARD; PRT; 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE A1 protein (40.8 kDa protein).  
 OS Beet curly top virus (BCTV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
 OX NCBI\_TaxID=10840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Pinnet M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 RT beet curly top virus."  
 RL EMBL J. 5:1761-1767(1986).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X04144; NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 DR ATP-binding. 222 229 ATP (POTENTIAL).  
 FT NP\_BIND  
 SQ SEQUENCE 358 AA; 40889 MW; 39AA5FE3C0B9C333 CRC64;

Query Match 59.0%; Score 213; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 4.6e-18;  
 Matches 39; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLWGEFQVDSRGSGCCTNSDAAAEALNASSKEEALQITAAIPEKYLFOFHNLSNL 60  
 DB 110 TLWGEFQIDGRSARGGQOTANDSYAKALNAGDVQSALNLKEOPKDYLVONHNIRSNL 169  
 OY 61 DRIFDKTPEP 70  
 DB 170 OKIFORPPDP 179

RESULT 12  
 VAL1\_TMOV STANDARD; PRT; 361 AA.  
 AC O06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein.  
 CN AL1.  
 OS Tomato mottle virus (isolate Florida) (TMOV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93107858; PubMed=1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 RT isolated from tomatoes in Florida."  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L14460; AAC32414.1; -.  
 DR PIR: J01870; J01870.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 DR ATP-binding. 222 229 ATP (BY SIMILARITY).  
 FT NP\_BIND  
 SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEBAC6950 CRC64;

Query Match 58.2%; Score 210; DB 1; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 1.1e-17;  
 Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLWGEFQVDSRGSGCCTNSDAAAEALNASSKEEALQITAAIPEKYLFOFHNLSNL 60  
 DB 110 TLWGEFQIDGRSARGGQOTANDSYAKALNAGDVQSALNLKEOPKDYLVONHNIRSNL 169  
 OY 61 DRIFDKTPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT; 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein (c1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RT with a single genomic component."  
 RL Virology 185:151-161(1991).

```

CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15656; CAA33688.1; -.
DR PIR: D40779; QCCVCL.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.1.
KW ATP-binding.
FT NP_BIND 219 226 ATP (POTENTIAL).
SQ SEQUENCE 357 AA; 40678 MW; 939AB6E1AB3B2A7 CRC64;

Query Match
Best Local Similarity 55.4%; Score 200; DB 1; Length 357;
Matches 39; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

OY 4 GEFQVDGRSARGCCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLQFHNLSNLDRT 63
DB 111 FGVSQIDGRSARGCOOSANDAYAEALNSGSEALNILEKAPKDYLLQFHNLSNLDRI 170
OY 64 F 64
DB 171 F 171

RESULT 14
VAL1_SLCV STANDARD; PRT; 347 AA.
ID VAL1_SLCV
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE AL1 protein.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082449; PubMed=1984668;
RA Lazarowitz S.G., Lazdins I.B.;
RT "Infectivity and complete nucleotide sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype."
RL Virology 180:58-69(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M38183; AAC32410.1; ALT_INIT.
DR PIR: C36785; QCCVSL.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.1.
KW ATP-binding.
FT NP_BIND 218 225 ATP (POTENTIAL).
SQ SEQUENCE 347 AA; 39110 MW; AFDAEBDE122110E CRC64;

```

```

Query Match
Best Local Similarity 35.5%; Score 128; DB 1; Length 347;
Matches 26; Conservative 13; Mismatches 23; Indels 4; Gaps 1;

OY 5 GEFQVDGRSARGCCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLQFHNLSNLDRT 64
DB 116 GQKVSQSG---GSKSKKDDYINAVNAGSAGBALIINAGDKPTIVYHNLSNLDRT 171
OY 65 DKTPPE 70
DB 172 QKPEPE 177

RESULT 15
Y175_HELPJ STANDARD; PRT; 299 AA.
ID Y175_HELPJ
AC Q92M07;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein JHP0161 precursor.
GN JHP0161.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyts G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -!- SIMILARITY: BELONGS TO THE PPTC/PARYULIN FAMILY OF ROTAMASES.
CC -----
CC STRONG, TO C.JEJUNI CBE2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001454; AAD05744.1; -.
DR HSSP: Q13526; 1PIN.
DR InterPro: IPR000297; Rotamase.
DR Pfam: PF00639; Rotamase; 1.
DR PROSITE: PS01096; PPTC_PPTASE_1; 1.
DR PROSITE: PS01098; PPTC_PPTASE_2; 1.
KW Hypothetical Protein; Isomerase; Rotamase; Signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 299 POTENTIAL.
FT DOMAIN 154 253 HYPOTHEETICAL PROTEIN JHP0161.
FT PPTC.
SQ SEQUENCE 299 AA; 34040 MW; 9C037B1D0D110143 CRC64;

Query Match
Best Local Similarity 17.0%; Score 61.5; DB 1; Length 299;
Matches 18; Conservative 12; Mismatches 13; Indels 21; Gaps 2;

OY 22 NDAAAEALNASSKEEALQIIAAAIPEKYLQFHNLSNLDRT 62
DB 93 NEAAEKLNQTPPEKAMEAVKQALVEFWAKKQAEVKKRIQIPEKMDQFY--NANKQD 150
OY 63 IFDK 66
DB 151 LFVK 154

```

Thu Oct 10 16:05:53 2002

us-09-289-346a-4.rsp

Page 7

Search completed: October 10, 2002, 02:09:51  
Job time: 1349 sec

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:43 ; Search time 69.63 Seconds  
(without alignments)  
96,600 Million cell updates/sec

Title: US-09-289-346A-4

Sequence: 1 TLVWGEFQVDGRSARGCGCT.....FOFHNLSNLDRIPTKPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	346	95.8	352	1	QOCVLI
2	249	69.0	361	1	QOCVPT
3	240	66.5	358	2	S07594
4	233	64.5	362	1	J01887
5	229	63.4	359	2	S39211
6	222	61.5	349	2	J02300
7	222	61.5	349	2	S31875
8	222	61.5	359	2	S22593
9	219	60.7	359	2	S39235
10	216	59.8	351	2	J02327
11	216	59.8	355	1	QOCVWL
12	213	59.0	385	1	S28360
13	210	58.2	358	1	J01870
14	200	55.4	357	1	QOCVCI
15	199	55.1	360	2	S59885
16	134	37.1	131	2	S45059
17	128	35.5	347	1	QOCVSI
18	72	19.9	587	2	J01419
19	68	18.8	714	2	C95382
20	63.5	17.6	584	2	T19061
21	62.5	17.3	1006	2	S74992
22	61.5	17.0	299	2	B71967
23	61.5	17.0	335	1	DBSSGF
24	61.5	17.0	481	2	A70091
25	61.5	17.0	1044	2	T43800
26	61.5	17.0	1203	2	T17415
27	61.5	17.0	1229	2	T48959
28	60.5	16.8	159	2	C71838
29	60.5	16.8	159	2	E64677

30	60.5	16.8	819	2	A13197
31	60	16.6	767	2	F71479
32	60	16.6	1008	2	T41244
33	59	16.3	160	2	G82060
34	59	16.3	201	2	G97513
35	59	16.3	201	2	AE2732
36	59	16.3	228	2	F90130
37	59	16.3	338	2	AG3582
38	59	16.3	1502	1	RGBYH1
39	58.5	16.2	297	1	MNVRV1
40	58.5	16.2	357	2	T02246
41	58.5	16.2	569	2	H87926
42	58.5	16.2	642	2	D90558
43	58.5	16.2	769	2	F81742
44	58.5	16.2	2135	2	T14602
45	58	16.1	345	2	AD3024

TonB-dependent rec  
hypothetical prote  
SCL4 protein homo  
hypothetical prote  
MDH ubiqnonone ox  
MDH ubiqnonone ox  
hypothetical prote  
iron(III)-binding  
CYC1/CYP3 transcri  
nonstructural prot  
hypothetical prote  
protein C17D12.7 l  
topoisomerase IV s  
conserved hypothet  
variant-specific s  
hypothetical prote

## ALIGNMENTS

RESULT 1

QOCVLI

ALI protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C>Date: 28-Aug-1985 #sequence-revision 28-Aug-1985 #text-change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A>Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.8%; Score 346; DB 1; Length 352;

Best local similarity 95.7%; Pred. No. 4e-33;

Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGCGCTSDNDAADALNASSKEALOTIAAIPKYLFOFHNLSNU 60

DB 111 TLVWGEFQVDGRSARGCGCTSDNDAADALNASSKEALOTIRKIPKYLFOFHNLSNU 170

QY 61 DRIFDKTPEP 70

DB 171 DRIFDKTPEP 180

RESULT 2

QOCVPT

ALI protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C>Date: 30-Jun-1992 #sequence-revision 30-Jun-1992 #text-change 16-Jun-2000

C:Accession: J00364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

A>Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: J00362; M01D:91311403

A:Accession: J00364

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-361 <COU>

A:Cross-references: GB:D00940; NID:q222458; PIDN:BAA00782.1; PID:q222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

QY	QY
1 TLVWGEQVODGRSARGGCQTSNDAAAEALNASSKEEALQITIAAAIPEKYLQFHNLNSNL	60
:        :	
:        :	

RESULT 7  
S31875  
All protein - pepper rizado amarillo virus  
C:Species: pepper rizado amarillo virus



Db 110 TAENGEEQIDRSRARGQQTANDSYAKALNAGDVQ\$ALNKEEQPDYVLQ\$HNHNSLN 169

Qy 61 DRFDKTPP 70  
: : : : :  
: : : : :  
: : : : :

Db 170 ERIFAKAP 179

RESULT 12

All protein - beet curly top virus  
 C:Species: beet curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Stanley, J.: Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 EMBO J. 5, 1761-1767, 1986  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top  
 A:Reference number: S28360  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GB: M24597; EMBL:X04144; NID:g210678; P1DN:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	59.0%;	Score 213;	DB 2;	Length 385;
Best Local Similarity	55.7%;	Pred. No. 2.4e-17;		
Matches 39;	Conservative 14;	Mismatches 17;	Indels 0;	Gaps 0;

[illegible]

RESULT 13

All protein - tomato mottle virus (isolate Florida)  
 C:Species: tomato mottle virus  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01870  
 R:Abouzid, A.M.; Polston, J.E.; Hiebert, E.  
 J. gen. Virol. 73, 3225-3229, 1992  
 A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from  
 A:Reference number: J01869; MUID:93107858  
 A:Accession: J01870  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <ABO>  
 A:Cross-references: GB:L14460  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match	58.28;	Score 210;	DB 1;	Length 358;
Best Local Similarity	54.38;	Pred. No. 5.1e-17;		
Matches	38;	Conservative	15;	Mismatches 17;
				Indels 0;
				Gaps 0;

[illegible]

RESULT 14

000VC1  
 A1: protein - tomato yellow leaf curl virus  
 N: Alternate names: C1 protein  
 C: Species: tomato yellow leaf curl virus  
 C: Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C: Accession: D40779  
 R: Nayot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A: Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single ORF  
 A: Reference number: A40779; MUID: 92024070  
 A: Accession: D40779  
 A: Status: translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-357 <NNAV>  
 A: Cross-references: GB:X15656; NID:g662204; PIDN:CAA33688.1; PID:g662207  
 C: Superfamily: tomato golden mosaic virus A1 protein

Query Match	55.48;	Score 200;	DB 1;	Length 357;
Best Local Similarity	63.98;	Pred. No. 7.8e-16;		
Matches 39;	Conservative 9;	Mismatches 13;	Indels 0;	Gaps 0;

[illegible]

QY	64	F	64
		-	
Db	171	F	171

## RESULT 1.5

replication-associated protein C1 - tomato yellow leaf curl virus  
 C.Species: tomato yellow leaf curl virus  
 C.Date: 14-Jan-1996 #sequence-revision 01-Mar-1996 #text-change 20-Sep-1999  
 C.Accession: S59885  
 R:Hong, Y., Harrison, B.D.  
 submitted to the EMBL Data Library, February 1995  
 A.Description: Nucleotide sequences from tomato leaf curl viruses from different countries  
 d.geminiviruses.  
 A.Reference number: S58346  
 A.Accession: S59885  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-360 <HON>  
 A.Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211  
 i:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	55.1%;	Score 199;	DB 2;	Length 360;
Best Local Similarity	57.6%;	Pred. No. 1e-15;		
Matches 38;	Conservative 12;	Mismatches 16;	Indels 0;	Gaps 0;

```
QY      4 WGEIQVDDGRSARGGCOTSNDAAAEALVAASSKEEALQIIIAAIPKEYLFPQHNLNSLDRI   63
          :|::||| ||| ::|||::|||::|||::|||::|||::|||::|||::|||
Db     113 FGVEQIDGRSARGGSOOSANDYAFAELINSSGAALDILREKAPKFVLQFHNLANNDRI 172
```

```
QY      64  FDKTP  69
        |  :  |
Db      173  FTPSAE 178
```

Search completed: October 10, 2002, 01:48:43  
Job time: 561 sec

	12	328	90.9	70	21	AAB18661	Mutant peptide der
	13	325	90.0	70	21	AAB18661	Mutant peptide der
	14	324	89.8	70	21	AAB18683	Mutant peptide der
	15	322	89.2	70	21	AAB18682	Mutant peptide der
	16	314	87.0	70	21	AAB18679	Mutant peptide der
	17	222	61.5	359	17	AAR88870	Sardinian tomato y
	18	222	61.5	359	17	AAR88871	Sardinian tomato y
	19	222	61.5	359	17	AAR88872	Sardinian tomato y
	20	219	60.7	353	18	AAW34338	Bean golden mosaic
	21	219	60.7	353	18	AAW34333	Bean golden mosaic
	22	219	60.7	353	18	AAW34332	Bean golden mosaic
	23	219	60.7	353	18	AAW34334	Bean golden mosaic
	24	219	60.7	353	18	AAW34335	Bean golden mosaic
	25	217	60.1	353	8	AAF70407	ORF 4 gene product
	26	214	59.3	361	18	AAW34336	Tomato mottle viru
	27	214	59.3	361	18	AAW34334	Tomato mottle viru
	28	214	59.3	361	18	AAW34335	Tomato mottle viru
	29	214	59.3	361	18	AAW34326	Tomato mottle viru
	30	203.5	56.4	361	8	AAV70562	Product of ORF 4 f
	31	200	55.4	357	18	AAW34339	Tomato yellow leaf
	32	200	55.4	357	18	AAW34330	Tomato yellow leaf
	33	200	55.4	357	18	AAW34331	Tomato yellow leaf
	34	198	54.8	362	19	AAW56495	Tobacco leaf curlI
	35	192	53.2	357	18	AAW34337	Tomato yellow leaf
	36	65.5	18.1	512	19	AAW68473	HIV-1 strain YBF30
	37	63.5	17.6	665	22	ABB70991	Drosophila melanog
	38	61.5	17.0	163	21	AAAG60400	Arabidopsis thalia
	39	61.5	17.0	299	18	AAW55450	H. pylori ORF Uzae
	40	61.5	17.0	1203	21	AAAY81275	Candida albicans C
	41	61.5	17.0	1203	21	AAAY84814	Amino acid sequenc
	42	61	16.9	483	22	AAAG31327	C glutamicum proteo
	43	60.5	16.8	131	18	AAW34323	Tomato mottle viru
	44	60.5	16.8	334	11	AAW03211	Amino acid sequenc
	45	60	16.6	214	20	AAV37449	Amino acid sequenc

ALIGNMENTS

RESULT	1	
AAAB18680	ID	AAAB18680 standard; peptide; 70 AA.
XX	AAAB18680;	
XX	22-JAN-2001	(first entry)
XX	Mutant peptide derived from amino acids 110-179 of Rep (All) protein.	
DE	Geminivirus; replication protein; Rep protein; All; transgenic plant;	
KM	ribosome binding region; resistance; geminivirus infection.	
XX	Synthetic.	
OS	Tomato golden mosaic virus.	
XX	Key	Location/Qualifiers
FH	Misc-difference 42	
FT	Misc-difference /note= "wild type residue replaced with Ala"	
FT	Misc-difference 43	
FT	/note= "wild type residue replaced with Ala"	
FT	Misc-difference 44	
FT	/note= "wild type residue replaced with Ala"	
PN	WO200054573-Al.	
XX	21-SEP-2000.	
PD	15-MAR-2000; 2000WO-US06759.	
PF	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX	(UYNC-) UNIV NORTH CAROLINA STATE.	

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 DR Transgenic plants with increased resistance to geminivirus infection  
 XX comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rd binding region  
 PT  
 XX  
 PS Claim 52; Page 43-44; 73pp; English.  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SO Sequence 70 AA;  
 Query Match 100.0%; Score 361; DB 21; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-39;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVYDGRSARCGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 60  
 DB 1 tlwgefyqdygrsarqgcqtsndaaeealnasskeeaqlilrekyllfqfhnlnsnl 60  
 QY 61 DRIEDKTPPEP 70  
 DB 61 driefdktppep 70  
 RESULT 2  
 AAB18677  
 ID AAB18677 standard; peptide; 70 AA.  
 AC AAB18677;  
 DT 22-JAN-2001 (first entry)  
 DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 KM Tomato golden mosaic virus.  
 OS  
 PN WO200054573-A1.  
 PD 21-SEP-2000.  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI: 2000-618851/59.  
 XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rd binding region  
 PT  
 XX  
 PS Disclosure; Page 18; 73pp; English.  
 CC The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX  
 SO Sequence 70 AA;  
 Query Match 95.8%; Score 346; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 1.7e-37;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVYDGRSARCGCOTSDNDAAEALNASSKEEALQIIAAIPEKYLFOFHNLNSNL 60  
 DB 1 tlwgefyqdygrsarqgcqtsndaaeealnasskeeaqlilrekyllfqfhnlnsnl 60  
 QY 61 DRIEDKTPPEP 70  
 DB 61 driefdktppep 70  
 RESULT 3  
 AAB18687  
 ID AAB18687 standard; peptide; 356 AA.  
 AC AAB18687;  
 DT 22-JAN-2001 (first entry)  
 DE Amino acid sequence of a geminivirus replication protein of TGMV.  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX Tomato golden mosaic virus.  
 OS  
 PN Key Location/Qualifiers  
 FT Misc-difference 354 /note="unspecified amino acid"  
 XX  
 PN WO200054573-A1.  
 PD 21-SEP-2000.  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI: 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 XX

PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT

PS Disclosure: Page 47-48; 73pp; English.

CC The present sequence represents a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.

XX Sequence 356 AA;

Query Match 95.8%; Score 346; DB 21; Length 356;  
 Best Local Similarity 95.7%; Pred. No. 1.4e-36;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TLWGEFQVNDGRSARGCOTSDNAEAALNASSKEEALQIIAAIPKYLFOFHNLNSL 60  
 Db 110 TLWGEFQVNDGRSARGCOTSDNAEAALNASSKEEALQIIKREKIPKYLIFQFNLNSL 169

OY 61 DRIFDKTPEP 70  
 Db 170 drifdktp 179

#### RESULT 4

AAB18685 ID AAB18685 standard; peptide: 70 AA.

XX AAB18685;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

OS Tomato golden mosaic virus.

XX Key Location/Qualifiers

FT Misc-difference 10 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

PI Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.

PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT

PS Claim 53; Page 46; 73pp; English.

CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.

XX Sequence 70 AA;

Query Match 93.6%; Score 338; DB 21; Length 70;  
 Best Local Similarity 94.3%; Pred. No. 1.8e-36;  
 Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TLWGEFQVNDGRSARGCOTSDNAEAALNASSKEEALQIIAAIPKYLFOFHNLNSL 60  
 Db 1 TLWGEFQVNDGRSARGCOTSDNAEAALNASSKEEALQIIKREKIPKYLIFQFNLNSL 169

OY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

#### RESULT 5

AAB18688 ID AAB18688 standard; peptide: 70 AA.

XX AAB18688;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

OS Tomato golden mosaic virus.

XX Key Location/Qualifiers

FT Misc-difference 19 /note= "wild type residue replaced with Ala"

FT Misc-difference 20 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

PI Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.

```

XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PS
PS Disclosure: Page 48; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;
Query Match 92.8%; Score 335; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 4, 5e-36;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDSARGGCGCTNDAAAEALNASSKEEALQIIIAAIPKRYLLFQFHNLSNL 60
DB 1 tlvwgefvydgsrargcgctndaaaealnaasskeeaqlilrekipexylifqfhnlnsl 60
QY 61 DRIFDKTPEP 70
DB 61 drifdktp 70
RESULT 6
AAB18692
ID AAB18692 standard; peptide: 70 AA.
AC
AC AAB18692;
XX
XX 22-JAN-2001 (first entry)
DT
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
DE
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
OS Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 66 /note= "wild type residue replaced with Ala"
FT Misc-difference 69 /note= "wild type residue replaced with Ala"
FT Misc-difference 69 /note= "wild type residue replaced with Ala"
FT
FT WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX
XX 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
PT

```

```

XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PS
PS Disclosure: Page 50; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;
Query Match 92.5%; Score 334; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 6, 1e-36;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDSARGGCGCTNDAAAEALNASSKEEALQIIIAAIPKRYLLFQFHNLSNL 60
DB 1 tlvwgefvydgsrargcgctndaaaealnaasskeeaqlilrekipexylifqfhnlnsl 60
QY 61 DRIFDKTPEP 70
DB 61 drifdktp 70
RESULT 7
AAB18684
ID AAB18684 standard; peptide: 70 AA.
AC
AC AAB18684;
XX
XX 22-JAN-2001 (first entry)
DT
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
DE
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
OS Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 7 /note= "wild type residue replaced with Ala"
FT Misc-difference 7 /note= "wild type residue replaced with Ala"
FT Misc-difference 8 /note= "wild type residue replaced with Ala"
FT
FT WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX
XX 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX
PT

```



XX PA Hanley-Bowdoin L, Orozco BM, Kong L;  
XX PI WPI: 2000-618851/59.  
XX DR  
XX PT Transgenic plants with increased resistance to geminivirus infection  
XX PT comprise a nucleic acid construct containing a nucleic acid sequence  
XX PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX PT  
XX PS Claim 52; Page 45; 73pp; English.  
XX  
XX CC The present sequence represents a mutant peptide, derived from a  
XX CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
XX CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
XX CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
XX CC protein are used to produce transgenic plants. The mutation in AL1 is  
XX CC present in a ribosome binding region, and expression of mutant AL1  
XX CC protein imparts increased resistance to geminivirus infection in the  
XX CC plant. Mutant AL1 proteins are useful for producing plants having  
XX CC increased resistance or reduced sensitivity to a geminivirus such as  
XX CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
XX CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
XX CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
XX CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
XX CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX SQ Sequence 70 AA;

Query Match 92.0%; Score 332; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 1.1e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGARSGCOTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSL 60  
Db 1 tlwgeaavdgrsargcgtsndaaealnaskeeaqlilrekipekylfifhnlsl 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

RESULT 8  
AAB18690  
ID AAB18690 standard; peptide: 70 AA.  
AC AAB18690;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX OS Synthetic.  
XX OS Tomato golden mosaic virus.  
XX  
XX FH Key Location/Qualifiers  
XX FH MISC-difference 27 /note= "wild type residue replaced with Ala"  
XX FT MISC-difference 30 /note= "wild type residue replaced with Ala"  
XX FT  
XX PN WO200054573-A1.  
XX PD 21-SEP-2000.  
XX PF 15-MAR-2000; 2000WO-US06759.  
XX PR 18-MAR-1999; 99US-0125004.  
XX PR 09-APR-1999; 99US-0289346.

XX PA (UNNC-) UNIV NORTH CAROLINA STATE.  
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX DR WPI: 2000-618851/59.  
XX PT Transgenic plants with increased resistance to geminivirus infection  
XX PT comprise a nucleic acid construct containing a nucleic acid sequence  
XX PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX PT  
XX PS Disclosure; Page 49; 73pp; English.  
XX  
XX CC The present sequence represents a mutant peptide, derived from a  
XX CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
XX CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
XX CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
XX CC protein are used to produce transgenic plants. The mutation in AL1 is  
XX CC present in a ribosome binding region, and expression of mutant AL1  
XX CC protein imparts increased resistance to geminivirus infection in the  
XX CC plant. Mutant AL1 proteins are useful for producing plants having  
XX CC increased resistance or reduced sensitivity to a geminivirus such as  
XX CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
XX CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
XX CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
XX CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
XX CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX SQ Sequence 70 AA;

Query Match 92.0%; Score 332; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 1.1e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGARSGCOTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSL 60  
Db 1 tlwgeaavdgrsargcgtsndaaealnaskeeaqlilrekipekylfifhnlsl 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

RESULT 9  
AAB18678  
ID AAB18678 standard; peptide: 70 AA.  
AC AAB18678;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX OS Synthetic.  
XX OS Tomato golden mosaic virus.  
XX  
XX FH Key Location/Qualifiers  
XX FH MISC-difference 12 /note= "wild type residue replaced with Ala"  
XX FT MISC-difference 13 /note= "wild type residue replaced with Ala"  
XX FT MISC-difference 15 /note= "wild type residue replaced with Ala"  
XX FT  
XX PN WO200054573-A1.  
XX PD 21-SEP-2000.  
XX

```

PF 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
XX Claim 53; Page 42-43; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 91.7%; Score 331; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1.5e-35;
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDGRSARGCGQTSNDAAAEALNASSKEEALQIIAAIPEKYLEFQFHNLSNL 60
DB 1 tlwgefgvdaaagsgqtsndaadaealnasskeeaqlirekipekylifqfhnlnsnl 60
QY 61 DRIFDKTPEP 70
DB 61 drifdktp 70
Db
RESULT 10
AAB18686
ID AAB18686 standard; peptide: 70 AA.
XX
XX AAB18686;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX
XX MISC-difference 24 /note= "wild type residue replaced with Leu"
XX
XX MISC-difference 25 /note= "wild type residue replaced with Leu"
XX
XX MISC-difference 26 /note= "wild type residue replaced with Leu"
XX
XX

```

```

PN WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
XX Claim 53; Page 46; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 91.7%; Score 331; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1.5e-35;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDGRSARGCGQTSNDAAAEALNASSKEEALQIIAAIPEKYLEFQFHNLSNL 60
DB 1 tlwgefgvdaagsgqtsndalllealnasskeeaqlirekipekylifqfhnlnsnl 60
QY 61 DRIFDKTPEP 70
DB 61 drifdktp 70
Db
RESULT 11
AAB18689
ID AAB18689 standard; peptide: 70 AA.
XX
XX AAB18689;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX
XX MISC-difference 22 /note= "wild type residue replaced with Ala"
XX
XX MISC-difference 23
XX

```

/note= "wild type residue replaced with Ala"

FT XX WO200054573-A1.  
 PN XX  
 PD XX  
 XX 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 DR WPI; 2000-618851/59.  
 XX  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rd binding region  
 PT  
 PS Disclosure; Page 48-49; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SO Sequence 70 AA;  
 91.4%; Score 330; DB 21; Length 70;  
 Query Match Best Local Similarity 92.9%; Pred. No. 2e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVYDGRSARGCQTSNDAAALNASSKEALQIIAAIPEKYLFOFHNLSNL 60  
 DB 1 TLVWGEFQVYDGRSARGCQTSNDAAALNASSKEALQIIAAIPEKYLFOFHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 DRIFDKTPEP 70  
 RESULT 12  
 AAB18691  
 ID AAB18691 standard; peptide: 70 AA.  
 XX  
 AC AAB18691;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 XX Tomato golden mosaic virus.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 34

/note= "wild type residue replaced with Ala"

FT FT Misc-difference 35  
 FT FT /note= "wild type residue replaced with Ala"  
 FT FT Misc-difference 36  
 FT FT /note= "wild type residue replaced with Ala"  
 PN XX WO200054573-A1.  
 PN XX  
 PD XX  
 XX 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 DR WPI; 2000-618851/59.  
 XX  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rd binding region  
 PT  
 PS Disclosure; Page 49; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SO Sequence 70 AA;  
 90.9%; Score 328; DB 21; Length 70;  
 Query Match Best Local Similarity 91.4%; Pred. No. 3.7e-35;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVYDGRSARGCQTSNDAAALNASSKEALQIIAAIPEKYLFOFHNLSNL 60  
 DB 1 TLVWGEFQVYDGRSARGCQTSNDAAALNASSKEALQIIAAIPEKYLFOFHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 DRIFDKTPEP 70  
 RESULT 13  
 AAB18681  
 ID AAB18681 standard; peptide: 70 AA.  
 XX  
 AC AAB18681;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.

OS	Tomato golden mosaic virus.	
XX	Key	Location/Qualifiers
FM	Misc-difference 47	/note= "wild type residue replaced with Ala"
FT	Misc-difference 48	/note= "wild type residue replaced with Ala"
FT	Misc-difference 49	/note= "wild type residue replaced with Ala"
FT	Misc-difference 49	/note= "wild type residue replaced with Ala"
PN	W0200054573-Al.	
XX	21-SEP-2000.	
PD	15-MAR-2000; 2000WO-US06759.	
XX	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
XX	Hanley-Bowdoin L, Orozco BM, Kong L.	
PI	WPI: 2000-618851/59.	
DR	Transgenic plants with increased resistance to geminivirus infection	
XX	comprise a nucleic acid construct containing a nucleic acid sequence	
PT	encoding a mutant Al1 protein with a mutation in the Rb binding region	
PT	-	
XX	Claim 52; Page 44; 73pp; English.	
PS	The present sequence represents a mutant peptide, derived from a	
XX	geminivirus replication (Rep) protein, also known as Al1. Al1 binds	
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded	
CC	DNA, and interacts with other viral and host proteins. Mutants of the Al1	
CC	protein are used to produce transgenic plants. The mutation in Al1 is	
CC	present in a ribosome binding region, and expression of mutant Al1	
CC	protein imparts increased resistance to geminivirus infection in the	
CC	plant. Mutant Al1 proteins are useful for producing plants having	
CC	increased resistance or reduced sensitivity to a geminivirus such as	
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl	
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian	
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic	
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper	
CC	virus, cotton leaf curl virus or beet curly top virus.	
CC	Sequence 70 AA;	
XX	SQ	
XX	Query Match 90.0%; Score 325; DB 21; Length 70;	
XX	Best Local Similarity 91.4%; Ped. No. 9.1e-35;	
XX	Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	1 TLVWGEFQVQGRARAGCGCTSDNDAAEALNASKKEALQIIAALPEKXYLFQFHNNLSNL 60	
Db	1 TLVWGEFQVQGRARAGCGCTSDNDAAEALNASKKEALQIIAALPEKXYLFQFHNNLSNL 60	
QY	61 DRIFDKTPRP 70	
Db	61 DRIFDKTPRP 70	
XX	RESULT 14	
ID	AAB18683 standard; peptide: 70 AA.	
XX	AAB18683;	
AC	22-JAN-2001 (first entry)	
XX	Mutant peptide derived from amino acids 110-179 of Rep (Al1) protein.	

KW	Geminiavirus; replication protein; Rep protein; AL1; transgenic plant;
KV	Geminiavirus; replication protein; Rep protein; AL1; transgenic plant;
XX	ribosome binding region; resistance; geminivirus infection.
OS	Synthetic.
XX	Tomato golden mosaic virus.
PH	Key
XX	Location/Qualifiers
FT	Misc-difference 59
FT	/note= "wild type residue replaced with Ala"
FT	Misc-difference 61
FT	/note= "wild type residue replaced with Ala"
FT	Misc-difference 62
FT	/note= "wild type residue replaced with Ala"
XX	WO200054573-A1.
XX	21-SEP-2000.
PD	
XX	15-MAR-2000; 2000WO-US06759.
PF	
XX	18-MAR-1999; 99US-0125004.
PR	09-APR-1999; 99US-0289346.
XX	(UYNC-) UNIV NORTH CAROLINA STATE.
PA	
XX	Hanley-Bowdoin L, Orozco BM, Kong L;
PI	WPI; 2000-618851/59.
XX	
DR	
XX	Transgenic plants with increased resistance to geminivirus infection
PT	comprise a nucleic acid construct containing a nucleic acid sequence
PT	encoding a mutant AL1 protein with a mutation in the Rb binding region
PT	-
XX	Claim 53; Page 45; 73pp; English.
XX	
XX	The present sequence represents a mutant peptide, derived from a
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC	protein are used to produce transgenic plants. The mutation in AL1 is
CC	present in a ribosome binding region, and expression of mutant AL1
CC	protein imparts increased resistance to geminivirus infection in the
CC	plant. Mutant AL1 proteins are useful for producing plants having
CC	increased resistance or reduced sensitivity to a geminivirus such as
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC	virus, cotton leaf curl virus or beet curly top virus.
XX	
SO	Sequence 70 AA;
	Query Match 89.8%; Score 324; DB 21; Length 70;
	Best Local Similarity 91.4%; Pred. No. 1,2e-34;
	Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY	1 TLWGEFQVNDGNSARCGCQTSMDAAEAALNMSKFEALQITAAIPKRYLFQFHNTNSMT 60
Db	1 tlvwgefqvndgnsarqcgctsndaaeaalnasskeeaqlirekipkylfqfhnlnsal 60
QY	61 DRIFDKRPP 70
Db	61 aalifdktrpp 70
RESULT 15	
AAB18682	
ID	AAB18682 standard; peptide; 70 AA.
XX	
XX	AAB18682;
XX	

Job time: 527 sec

DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 52 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 54 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 PT Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rd binding region  
 PT  
 XX  
 PS Claim 53; Page 44-45; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1 is  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 XX

Query Match 89.2%; Score 322; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 2.2e-34;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 TLVWGFQVGVGRSARGCCGCTSNDAAEALNASSKEEALQIIAAIPKYLPOPHNLNSNC 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 tlvwgfqvgvgrsargcgctsndaaealnasskeaalqilirekpekylfafaalnsnl 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

OY 61 DRIFDKTPEP 70  
 ||||||||||||  
 Db 61 drifdktp 70



Thu Oct 10 16:05:59 2002

us-09-289-346a-5.ispt

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds  
(without alignments)  
102.295 Million cell updates/sec

Title: US-09-289-346a-5  
Perfect score: 359  
Sequence: 1 TLWGEFQYDGRSARGCCT.....PQFHNLSNLDRIEDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

- 1: SP\_ARCHAEA:\*
- 2: SP\_BACTERIA:\*
- 3: SP\_FUNGI:\*
- 4: SP\_HUMAN:\*
- 5: SP\_INVERTEBRATE:\*
- 6: SP\_MAMMAL:\*
- 7: SP\_MHC:\*
- 8: SP\_ORGANELLE:\*
- 9: SP\_PHAGE:\*
- 10: SP\_PLANT:\*
- 11: SP\_RODENT:\*
- 12: SP\_VIRUS:\*
- 13: SP\_VERTEBRATE:\*
- 14: SP\_UNCLASSIFIED:\*
- 15: SP\_VIRUS:\*
- 16: SP\_BACTERIAP:\*
- 17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	298	83.0	351	12 Q91R10	Q91R10 tomato seve
2	296	82.5	352	12 Q9E000	Q9E000 tomato rugo
3	287	79.9	226	12 Q9WHF6	Q9WHF6 tomato mild
4	277	77.2	361	12 Q67574	Q67574 bean golden
5	275	76.6	225	12 Q90DB1	Q90DB1 cowpea gold
6	275	76.6	226	12 Q09727	Q09727 leonurus mo
7	271	75.5	185	12 Q98693	Q98693 sida golden
8	271	75.5	314	12 Q9ELT8	Q9ELT8 sweet potat
9	271	75.5	364	12 Q9Q5S5	Q9Q5S5 sweet potat
10	264	73.5	149	12 P88975	P88975 macropyllin
11	264	73.5	190	12 Q9Z089	Q9Z089 tobacco lea
12	264	73.5	190	12 Q9Z084	Q9Z084 tobacco lea
13	264	73.5	233	12 Q9YLA4	Q9YLA4 macropyllin
14	263	73.3	190	12 Q9W827	Q9W827 tobacco lea
15	263	73.3	208	12 Q9Z0C4	Q9Z0C4 tobacco lea
16	263	73.3	359	12 Q91M88	Q91M88 tobacco lea

RESULT	ID	PRELIMINARY:	PRT:	351 AA.	ALIGNMENTS
17	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
18	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
19	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
20	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
21	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
22	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
23	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
24	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
25	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
26	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
27	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
28	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
29	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
30	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
31	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
32	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
33	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
34	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
35	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
36	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
37	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
38	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
39	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
40	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
41	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
42	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
43	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
44	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
45	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10

Query Match: 83.0%, Pred. No. 8.5e-26; Length 351;  
Best Local Similarity: 80.0%; Mismatches: 8; Indels: 0; Gaps: 0;  
Matches: 56; Conservative: 6; Mismatches: 8; Indels: 0; Gaps: 0;  
DB: 111 TLWGEFQYDGRSARGCCTNDAAAEALNMSKREALQIREKIPPAALFOFHNLSNL 60  
QY: 1 TLWGEFQYDGRSARGCCTNDAAAEALNMSKREALQIREKIPPAALFOFHNLSNL 60  
DB: 111 TLWGEFQYDGRSARGCCTNDAAAEALNMSKREALQIREKIPPAALFOFHNLSNL 170  
QY: 61 DRIFDKTPPEP 70  
DB: 171 DRIFDKTPPEP 180  
RESULT: 2  
Q9E000  
ID: Q9E000  
PRELIMINARY: PRT: 352 AA.

AC 09E000;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN.  
 GN AC1.  
 OS Tomato rugose mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=134599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fernandes J.J., Fontes F.P.B., Brommonschenkel S.H., Carvalho M.G.,  
 RA Zambolim E.M., Zerbini F.M.,  
 RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
 (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,  
 RT Minas Gerais, Brazil.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF291705; AAC15546.1; -;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Geminl\_AL1.  
 SO SEQUENCE 352 AA; 40012 MW; 47CD5838E24D613 CRC64;

Query Match 82.5%; Score 296; DB 12; Length 352;  
 Best Local Similarity 80.0%; Pred. No. 1.4e-25;  
 Matches 56; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSRGSGCOTNDAAAEALNASSKEEALQIIREKTPAALFOFHNLNSNL 60  
 DB 111 TLVWGEFQVDSRGSGCOTNDAAAEALNASSKEEALQIIREKTPAALFOFHNLNSNL 170  
 OY 61 DRIFDKTPP 70  
 DB 171 DRIFAKAPP 180

RESULT 3  
 O9WHF6 PRELIMINARY; PRT; 226 AA.  
 AC O9WHF6;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Tomato mild mottle geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
 OX NCBI\_TaxID=92943;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-HN96-H5KW;  
 RA Nahla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
 RA Maxwell D.P.;  
 RT "Molecular characterization and DNA-based detection methods for  
 RT vegetable-infecting geminiviruses in Central America.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; A131071; AAD33471.1; -;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Geminl\_AL1.  
 FT NON\_TER 226  
 SO SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 79.9%; Score 287; DB 12; Length 226;  
 Best Local Similarity 77.1%; Pred. No. 9.2e-25;  
 Matches 54; Conservative 9; Mismatches 7; Indels 0; Gaps 0;  
 OY 1 TLVWGEFQVDSRGSGCOTNDAAAEALNASSKEEALQIIREKTPAALFOFHNLNSNL 60  
 DB 111 TLVWGEFQVDSRGSGCOTNDAAAEALNASSKEEALQIIREKTPAALFOFHNLNSNL 170

DB 111 TLVWGEFQVDSRGSGCOTNDAAAEALNASSKEEALQIIREKTPAALFOFHNLNSNL 170  
 OY 61 DRIFDKTPP 70  
 DB 171 DRIFAKAPP 180

RESULT 4  
 ID 067574 PRELIMINARY; PRT; 361 AA.  
 AC 067574;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN AL1.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 RA Maxwell D.P., Russell D.R.;  
 RT "Cloning of the complete DNA genomes of four bean-infecting  
 RT geminiviruses and determining their infectivity by electric discharge  
 RT particle acceleration.";  
 RL Phytopathology 81:980-985(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 RA Morales F.J., Maxwell D.P.;  
 RT "Differentiation of bean-infecting geminiviruses by nucleic acid  
 RT hybridization probes and aspects of bean golden mosaic in Brazil.";  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
 RT "Genetic diversity in geminiviruses causing bean golden mosaic  
 RT disease: The nucleotide sequence of the infectious cloned DNA  
 RT components of a Brazilian isolate of bean golden mosaic geminivirus.";  
 DR EMBL; M88686; AAA46312.1; -;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Geminl\_AL1.  
 SO SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.2%; Score 277; DB 12; Length 361;  
 Best Local Similarity 77.6%; Pred. No. 2.2e-23;  
 Matches 52; Conservative 8; Mismatches 7; Indels 0; Gaps 0;  
 OY 4 WGEFQVDSRGSGCOTNDAAAEALNASSKEEALQIIREKTPAALFOFHNLNSNDRI 63  
 DB 113 WGEFQVDSRGSGCOTNDAAAEALNASSKEEALQIIREKTPAALFOFHNLNSNDRI 172  
 OY 64 FDKTPP 70  
 DB 173 FTRKAPP 179

RESULT 5  
 ID 090DB1 PRELIMINARY; PRT; 225 AA.  
 AC 090DB1;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.



OX NCBI\_TaxID=69263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGMV-BR;  
 RA Faria J.C.;  
 RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from Brazil."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF188708; AAF06318.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 225 225  
 FT SEQUENCE 225 AA; 25766 MW; 1089CB6BD8D15B5D CRC64;

Query Match 76.6%; Score 275; DB 12; Length 225;  
 Best Local Similarity 77.6%; Pred. No. 2.1e-23;  
 Matches 52; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 4 WGEFQVVDGSRAGCGCOTSDNAAEALNASSKEEALQIIRKIPAAALFOFHNLNSLDR 63  
 DB 113 WGFQIDGSRAGGQOTINDAAEALNASSKEEAMQIKELPEKFLFOYHNLNSLDR 172  
 OY 64 FDKTPPEP 70  
 DB 173 FKKPEP 179

RESULT 6  
 009727 PRELIMINARY; PRT: 226 AA.  
 ID 009727;  
 AC 009727;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonurus mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=58177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LEMV-BRAZIL 1;  
 RA Faria J.C.; Maxwell D.P.;  
 RT "Variability in geminivirus associated with Phaseolus vulgaris in Brazil."  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U92532; AAB51157.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 226 226  
 FT SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 76.6%; Score 275; DB 12; Length 226;  
 Best Local Similarity 77.1%; Pred. No. 2.1e-23;  
 Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVWGEFQVVDGSRAGCGCOTSDNAAEALNASSKEEALQIIRKIPAAALFOFHNLNSL 60  
 DB 111 TLVWGEFQVVDGSRAGGQOTVNDAAEALNAPDKRTALQIKELPEKFLFOFHNLNSL 170  
 OY 61 DRIFDKTPPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 7  
 098693

ID 098693 PRELIMINARY; PRT: 185 AA.  
 AC 098693;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN AC1.  
 OS sida golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51034;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
 RT "Genetic Diversity among geminiviruses associated with the weed species Sida spp., Macroptilium lathyroides, and Wissadula amplissima from Jamaica."  
 RL Plant Dis. 81:1251-1258(1997).  
 DR EMBL: U67926; AAB97865.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 185 185  
 FT SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 75.5%; Score 271; DB 12; Length 185;  
 Best Local Similarity 71.4%; Pred. No. 4.8e-23;  
 Matches 50; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

OY 1 TLVWGEFQVVDGSRAGCGCOTSDNAAEALNASSKEEALQIIRKIPAAALFOFHNLNSL 60  
 DB 89 TLVWGFQIDGSRAGGQOTANDAAEALNASTCKEDALKIIRKIPERYLFOYHNLSSNI 148  
 OY 61 DRIFDKTPPEP 70  
 DB 149 DRIFSKPEP 158

RESULT 8  
 098693 PRELIMINARY; PRT: 314 AA.  
 ID 098693;  
 AC 098693;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United States."  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF288227; AAG01006.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 75.5%; Score 271; DB 12; Length 314;  
 Best Local Similarity 79.4%; Pred. No. 8.8e-23;



DE AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA5;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001318; BAA34039.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 1 1  
 FT 190 190  
 SQ SEQUENCE 190 AA: 21444 MW: AACIC2943EEF01AD CRC64;

Query Match 73.5%; Score 264; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 3.1e-22;  
 Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;

OY 1 TLWGFEVDGRSARGCGCOTNDAAAEALNASSKEEALOTIRKIPPAALFOFHNLSNL 60  
 DB 85 TLEMGTFQIDGRSARGCGCOTNDAAAEALNASSKEEALOTIRKIPFOYHNLSNL 144  
 OY 61 DRI-----FKTPE 69  
 DB 145 DRIFAPPLEVFCPSNSSFDVPE 169

RESULT 13  
 O9YL44  
 ID O9YL44 PRELIMINARY; PRT; 233 AA.  
 AC O9YL44;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Macroplidium golden mosaic geminivirus.  
 OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.;  
 RT "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses from Jamaica."  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E., McLaughlin W.A., Maxwell D.P.;  
 RT "Molecular characterization of two distinct geminiviruses infecting M. lachryoides from Jamaica."  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF098940; AAD17850.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 233 233  
 SQ SEQUENCE 233 AA: 26356 MW: AA490AF4D2166A02 CRC64;

Query Match 73.5%; Score 264; DB 12; Length 233;  
 Best Local Similarity 70.0%; Pred. No. 3.9e-22;  
 Matches 49; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

OY 1 TLWGFEVDGRSARGCGCOTNDAAAEALNASSKEEALOTIRKIPPAALFOFHNLSNL 60  
 DB 110 TLEMGTFQIDGRSARGCGCOTNDAAAEALNASSKEEALOTIRKIPFOYHNLSNL 169  
 OY 61 DRIFDKTPE 70  
 DB 170 DRIFMDPEP 179

RESULT 14  
 O9W827  
 ID O9W827 PRELIMINARY; PRT; 190 AA.  
 AC O9W827;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORAL-1, PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GORAL;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001303; BAA34010.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 1 1  
 FT 190 190  
 SQ SEQUENCE 190 AA: 21444 MW: 93C3742A8EBDB7EB CRC64;

Query Match 73.3%; Score 263; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 4e-22;  
 Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;

OY 1 TLWGFEVDGRSARGCGCOTNDAAAEALNASSKEEALOTIRKIPPAALFOFHNLSNL 60  
 DB 85 TLEMGTFQIDGRSARGCGCOTNDAAAEALNASSKEEALOTIRKIPFOYHNLSNL 144  
 OY 61 DRI-----FKTPE 69  
 DB 145 DRIFAPPLEVFCPSNSSFDVPE 169

RESULT 15  
 O9Z0C4  
 ID O9Z0C4 PRELIMINARY; PRT; 208 AA.  
 AC O9Z0C4;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ABURA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001294; BAA33992.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:51 ; Search time 34.62 Seconds  
(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346a-5

Perfect score: 359  
Sequence: 1 TLVWGEFQVDRSARAGCGCT.....FOFHNLNSNDRIEFDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	343	95.5	352	1 VAL1_TGMV	P03567 tomato gold
2	247	68.8	361	1 VAL1_PYMV	P27258 potato yell
3	243	67.7	358	1 VAL1_CLVX	P14982 cassava lat
4	243	67.7	358	1 VAL1_CLVN	P14972 cassava lat
5	232	64.6	362	1 VAL1_TYCA	P36279 tomato yell
6	226	63.0	359	1 VAL1_TYCU	P36269 tomato yell
7	224	62.4	349	1 VAL1_PHUV	P06923 pepper huas
8	219	61.0	359	1 VAL1_TYCM	P27260 tomato yell
9	216	60.2	353	1 VAL1_BGMV	P05175 bean golden
10	216	60.2	361	1 VAL1_TMOV	P06657 tomato molt
11	215	59.9	355	1 VAL1_ABMV	P21947 abutilon mo
12	212	59.1	358	1 VAL1_BCTV	P14991 beet curly
13	203	56.5	357	1 VAL1_TYCV	P27259 tomato yell
14	118	32.9	347	1 VAL1_SLVC	P29048 squash leaf
15	65	18.1	630	1 VE1_HPV66	O80957 human papil
16	63.5	17.7	1713	1 LMA3_HUMAN	O16787 human papil
17	62.5	17.4	428	1 GBAL_CANAL	P28868 candida alb
18	61	17.0	431	1 WFL_HPV30	O05112 human papil
19	60.5	16.9	447	1 TBL2_HUMAN	O9Y4P3 homo sapien
20	60.5	16.9	1070	1 P11B_HUMAN	P42338 homo sapien
21	59.5	16.6	706	1 Y006_RICPR	O926d6 ticketasia
22	59.5	16.6	1053	1 UBA3_WHEAT	P31252 triticum ae
23	59.5	16.6	1610	1 CCAD_MESAU	O99244 mesocricetu
24	59.5	16.6	2161	1 CCAD_HUMAN	O01668 homo sapien
25	59.5	16.6	2203	1 CCAD_RAT	P27732 rattus norv
26	59	16.4	703	1 GYS2_RAT	P17625 rattus norv
27	59	16.4	1852	1 CCAS_CYPRA	P22316 cyprinus ca
28	58	16.2	387	1 Y4PE_RHISN	P55615 rhizobium s
29	58	16.2	1127	1 Y855_TREPA	O83827 treponema p
30	58	16.2	1608	1 HLXA_SERNA	P15320 serratalia ma
31	57.5	16.0	367	1 LHXA_MOUSE	P53776 mus musculu
32	57.5	16.0	511	1 HUPH_VIBCH	O9Ksq4 vibrio chol
33	57	15.9	355	1 CRFB_RHOSH	P54905 rhodobacter

34	57	15.9	428	1 FLT1_MOUSE	O08917 mus musculu
35	57	15.9	432	1 PROA_DETRA	O9ctd9 deinococcus
36	57	15.9	524	1 MPPA_MOUSE	O9qcd1 mus musculu
37	57	15.9	524	1 MPPA_MOUSE	P20069 rattus norv
38	57	15.9	874	1 SLAP_BACLI	P49052 bacillus li
39	56.5	15.7	1070	1 P11B_RAT	O92110 rattus norv
40	56	15.6	129	1 RK12_PORPU	P51339 porphyra pu
41	56	15.6	329	1 TC3A_CAREL	P34257 caenorhabdi
42	56	15.6	351	1 PSTS_RHITO	O98f12 rhizobium l
43	56	15.6	427	1 FLT1_HUMAN	O75955 homo sapien
44	56	15.6	428	1 FLT1_RAT	O921e1 rattus norv
45	55.5	15.5	266	1 EYFB_MYCLE	O33095 mycobacteri

## ALIGNMENTS

RESULT	1	STANDARD	PRT	352 AA
VAL1_TGMV				
ID VAL1_TGMV				
AC P03567:				
DT 21-JUL-1986 (Rel. 01, Created)				
DT 21-JUL-1986 (Rel. 01, Last sequence update)				
DT 01-JUN-1994 (Rel. 29, Last annotation update)				
DE AL1 protein.				
GN ACT.				
OS Tomato golden mosaic virus (TGMV).				
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.				
OX NCBI_TaxID:10831;				
RN [1]				
RP SEQUENCE FROM N.A.				
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;				
RT "Complete nucleotide sequence of the infectious cloned DNA components				
RT of tomato golden mosaic virus: potential coding regions and regulatory				
RT sequences."				
RL EMBL J.3:2197-2205(1984).				
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC the European Bioinformatics Institute. There are no restrictions on its				
CC use by non-profit institutions as long as its content is in no way				
CC modified and this statement is not removed. Usage by and for commercial				
CC entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC				
DR EMBL: K02029; -; NOT_ANNOTATED_CDS.				
DR PIR: A04170; QOCVLI.				
DR InterPro: IPR001191; Gemin1_AL1.				
DR Pfam: PF00799; Gemin1_AL1; 1.				
DR PRINTS: PR000227; GEMCOATL1.				
DR PRODOM: PD000736; Gemin1_AL1; 1.				
KW ATP-binding; 223				
FT NE-BIND				
FT SEQUENCE 352 AA; 40332 MW; C33C938B644B4M4 CRC64;				

Query Match 95.5%; Score 343; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 1.3e-31;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	TLVWGEFQVDRSARAGCGCTSNDAAEALNASSKEFEALDIIEKIPPAALFOFHNLNSNL 60
DB	111	TLVWGEFQVDRSARAGCGCTSNDAAEALNASSKEFEALDIIEKIPPEYTLQFHNLNSNL 170
QY	61	DRIFDKTPPEP 70
DB	171	DRIFDKTPPEP 180

RESULT 2  
VAL1\_PYMV STANDARD: PRT: 361 AA.  
ID VAL1\_PYMV

```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DR 01-AUG-1992 (Rel. 23, Last sequence update)
DE 01-AUG-1992 (Rel. 23, Last annotation update)
OS AL1 protein.
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311403; PubMed=1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RL J. Gen. Virol. 72:1515-1520(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00940; BAA00782.1; -
DR PIR; J00364; OGCVP.
DR InterPro; IPR001191; Gemin1_AL1.
DR Pfam; PF00799; Gemin1_AL1.
DR PRINTS; PR00227; GEMCOAT1.1.
DR Prodom; PD000736; Gemin1_AL1.1.
KW ATP-binding.
FT NP_BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match
Best Local Similarity 68.8%; Score 247; DB 1; Length 361;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVGEFQVDSRGSGCQTSNDAAAEALNASSKEALQIIREKIPPAALFOFHNLNSNL 60
ID VAL1_CLV1 STANDARD; PRT; 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
DE AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Gay M.R.;
RT "Nucleotide sequence of cassava latent virus DNA."
RL Nature 301:260-262(1983).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02057; -; NOT ANNOTATED_CDS.
DR InterPro; IPR001191; Gemin1_AL1.
DR Pfam; PF00799; Gemin1_AL1.
DR PRINTS; PR00227; GEMCOAT1.1.
DR Prodom; PD000736; Gemin1_AL1.1.
KW ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40346 MW; ED173E753BE92D69 CRC64;

Query Match
Best Local Similarity 67.7%; Score 243; DB 1; Length 358;
Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLVGEFQVDSRGSGCQTSNDAAAEALNASSKEALQIIREKIPPAALFOFHNLNSNL 60
ID VAL1_CLV1 STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
DE AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RL African cassava mosaic virus (Nigerian strain).";
RL Nucleic Acids Res. 18:197-198(1990).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17095; CAA34953.1; -
DR PIR; S07594; S07594.
DR InterPro; IPR001191; Gemin1_AL1.
DR Pfam; PF00799; Gemin1_AL1.
DR PRINTS; PR00227; GEMCOAT1.1.
DR Prodom; PD000736; Gemin1_AL1.1.
KW ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40435 MW; 1DB16B8CB2D52C CRC64;

Query Match
Best Local Similarity 67.7%; Score 243; DB 1; Length 358;
Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLVGEFQVDSRGSGCQTSNDAAAEALNASSKEALQIIREKIPPAALFOFHNLNSNL 60
ID VAL1_CLV1 STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
DE AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RL African cassava mosaic virus (Nigerian strain).";
RL Nucleic Acids Res. 18:197-198(1990).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

QY 61 DRIFDKTPEP 70  
 DB 169 DRIFQEPAP 178

RESULT 5  
 VAL1\_TYLCU

ID VAL1\_TYLCU STANDARD: PRT: 362 AA.  
 AC P36279;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE AL1 protein (CI protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36447;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9339778; PubMed=8423446;  
 RA Dry I.B., Rigden J.E., Krahe L.R., Mullineaux P.M., Rezaian M.A.:  
 RT "Nucleotide sequence and genome organization of tomato leaf curl  
 geminivirus."  
 RL J. Gen. Virol. 74:147-151(1993).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 DR PIR: J01887; J01887.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND  
 SQ SEQUENCE 362 AA: 41197 MW: 343E718484704098 CRC64:

Query Match 64.6%; Score 232; DB 1; Length 362;  
 Best Local Similarity 55.3%; Pred. No. 5.3e-19;  
 Matches 47; Conservative 8; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLWGEFQVDRGSRAGCGCOTSDNAAEALNASSKEALDIIRKIPAAALFQPHNLNSML 60  
 DB 110 TLWGEFQIDGRSARGGQQTANDAYAKAINAGSKSEALDIKELAPROYILHFHNSMLD 169  
 QY 61 DRI-----FDKTPPE 69  
 DB 170 DRIFTPLEYYVSPFLSSSDRPE 194

RESULT 6  
 VAL1\_TYLCU

ID VAL1\_TYLCU STANDARD: PRT: 359 AA.  
 AC P38609;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE AL1 protein (CI protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=37139;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94256836; PubMed=8198442;  
 RA Morris E., Hidalgo E., Accotto G., Moriones E.:  
 RT "High similarity among the tomato yellow leaf curl virus isolates  
 from the west Mediterranean basin: the nucleotide sequence of an  
 infectious clone from Spain."  
 RL Arch. Virol. 135:165-170(1994).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z25751; CAA81026.1; -  
 DR PIR: S39211; S39211.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND  
 SQ SEQUENCE 359 AA: 41065 MW: 2D170A51EF80A3BC CRC64:

Query Match 63.0%; Score 226; DB 1; Length 359;  
 Best Local Similarity 60.9%; Pred. No. 2.5e-18;  
 Matches 42; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 2 LWMGEFQVDRGSRAGCGCOTSDNAAEALNASSKEALDIIRKIPAAALFQPHNLNSMLD 61  
 DB 111 LWMGEFQIDGRSARGGQQTANDAYAKAINAGSKSEALDIKELAPROYILHFHNSMLD 170  
 QY 62 RIFDKTPEP 70  
 DB 171 RVFQVPPAP 179

RESULT 7  
 VAL1\_PHVU

ID VAL1\_PHVU STANDARD: PRT: 349 AA.  
 AC Q06923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN AL1.  
 OS Pepper huasteco virus (PHV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=28349;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94015007; PubMed=8409944;  
 RA Torres-Pacheco I., Garzon-Riznado J.A., Herrera-Estrella L.,  
 RA Rivera-Bustamante R.F.:  
 RT "Complete nucleotide sequence of pepper huasteco virus: analysis and  
 comparison with bipartite geminiviruses."  
 RL J. Gen. Virol. 74:2225-2231(1993).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X70418; CAA49856.1; -  
 DR PIR: S31875; S31875.  
 DR PIR: J02300; J02300.  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND  
 SQ SEQUENCE 349 AA: 39722 MW: D5F4E76CD56370F4 CRC64:

Query Match 62.4%; Score 224; DB 1; Length 349;

```

Best Local Similarity 61.4%; Pred. No. 41e-18;
Matches 43; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFOVDGSRAGCGCOTSDNDAEAALNASSKEEALQIREKIPAAALFQFHNLSNLD 60
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLWGEFOVDGSRAGCGCOTSDNDAEAALNASSKEEALQIREKIPAAALFQFHNLSNLD 169
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIEDKTPPEP 70
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 NRIFQTPPEP 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
VAL1_TYLCM STANDARD; PRT: 359 AA.
ID VAL1_TYLCM STANDARD; PRT: 359 AA.
AC P27260;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE A11 protein (C1 protein).
CN 1.
OS Tomato yellow leaf curl virus (strain Marmagne) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92107660; PubMed=1840676;
RA Kheyr-Pour A., Bendahmane M., Matzelt V., Accotto G.P., Crespi S.,
RT Gronenborn B.;
RT "Tomato yellow leaf curl virus from Sardinia is a
RL Whitefly-transmitted monopartite geminivirus.";
RL Nucleic Acids Res. 19:6763-6769(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: X61153; CAA43466.1; -
CC PIR: S22593; S22593.
CC InterPro: IPR001191; Gemin1_A11.
CC Pfam: PF00799; Gemin1_A11; 1.
CC PRINTS: PR00227; GEMCOAT1.1.
CC ProDom: PD000736; Gemin1_A11; 1.
CC ATP-binding.
CC NP_BIND 220 227 ATP (POTENTIAL).
CC FT SEQUENCE 359 AA; 40733 MW; 9717B4A07C93EFA7 CRC64;
CC
Query Match 61.0%; Score 219; DB 1; Length 359;
Best Local Similarity 58.0%; Pred. No. 1.6e-17;
Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVMGEFOVDGSRAGCGCOTSDNDAEAALNASSKEEALQIREKIPAAALFQFHNLSNLD 61
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 LVMGEFOVDGSRAGCGCOTSDNDAEAALNASSKEEALQIREKIPAAALFQFHNLSNLD 170
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 62 DRIEDKTPPEP 70
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 171 KVFQVPAP 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
VAL1_BGMV STANDARD; PRT: 353 AA.
ID VAL1_BGMV STANDARD; PRT: 353 AA.
AC P05175;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)

```

```

DE A11 protein (40.2 kDa protein).
CN A11.
OS Bean golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10839;
RN [1]
RP SEQUENCE FROM N.A.
RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene
RL regulation in geminiviruses.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: M10070; AAA46318.1; -
CC InterPro: IPR001191; Gemin1_A11.
CC Pfam: PF00799; Gemin1_A11; 1.
CC PRINTS: PR00227; GEMCOAT1.1.
CC ProDom: PD000736; Gemin1_A11; 1.
CC ATP-binding.
CC NP_BIND 222 229 ATP (POTENTIAL).
CC FT SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;
CC
Query Match 60.2%; Score 216; DB 1; Length 353;
Best Local Similarity 60.0%; Pred. No. 3.3e-17;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFOVDGSRAGCGCOTSDNDAEAALNASSKEEALQIREKIPAAALFQFHNLSNLD 60
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLWGEFOVDGSRAGCGCOTSDNDAEAALNASSKEEALQIREKIPAAALFQFHNLSNLD 169
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIEDKTPPEP 70
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 ERLEFVPEP 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
VAL1_TMOV STANDARD; PRT: 361 AA.
ID VAL1_TMOV STANDARD; PRT: 361 AA.
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE A11 protein.
CN A11.
OS Tomato mottle virus (isolate Florida) (TMOV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=36449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107658; PubMed=1469361;
RA Abouzaid A.M., Polston J.E., Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RL isolated from tomatoes in Florida.";
RL J. Gen. Virol. 73:3225-3229(1992).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC

```



DR EMBL: L14460; AAC2414.1; -  
 DR PIR: J01870; J01870.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP-BIND 222 229 ATP (BY SIMILARITY).  
 SO SEQUENCE 361 AA; 40516 MW; 8138B6CCEAC6950 CRC64;

Query Match

Best Local Similarity 60.2%; Score 216; DB 1; Length 361;  
 Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARSGCOTSNDAAEALNASSKEEALQIREKIPAAALFQFHNLNSNL 60  
 DB 110 TIEWGDFQIDGRSARSGQOSANDSYAKALNAGVSALVLRDEQPKDFVLQNHNRSL 169  
 QY 61 DRIFDKTPPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 11

VAL1\_ABMVW STANDARD: PRT; 355 AA.  
 AC P21947;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein.  
 GN A1.  
 OS Abutilon mosaic virus (isolate West India).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91020984; PubMed-2219703;  
 RA Fritschmuth T., Zimmat G., Jeske H.;  
 RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
 RL as well as eukaryotic features."  
 RL Virology 178:461-468(1990).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X15983; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; QOCVWL.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP-BIND 221 228 ATP (POTENTIAL).  
 SO SEQUENCE 355 AA; 40257 MW; 16A2CABA63251E95 CRC64;

Query Match 59.9%; Score 215; DB 1; Length 355;  
 Best Local Similarity 58.6%; Pred. No. 4.4e-17;  
 Matches 41; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARSGCOTSNDAAEALNASSKEEALQIREKIPAAALFQFHNLNSNL 60  
 DB 110 TIEWGDFQIDGRSARSGQOTANDSYAKALNAGVSALVLRDEQPKDFVLQNHNRSL 169  
 QY 61 DRIFDKTPPEP 70

DB 170 ERIFAKAPEP 179

RESULT 12

VAL1\_BCTV STANDARD: PRT; 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE A1 protein (40.8 kDa protein).  
 OS Beet curly top virus (BCTV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
 OX NCBI\_TaxID=10840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 RT beet curly top virus."  
 RL EMBO J. 5:1761-1767(1986).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X04144; -; NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP-BIND 222 229 ATP (POTENTIAL).  
 SO SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 59.1%; Score 212; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 9.6e-17;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARSGCOTSNDAAEALNASSKEEALQIREKIPAAALFQFHNLNSNL 60  
 DB 110 TIEWGDFQIDGRSARSGQOTANDSYAKALNAGVSALVLRDEQPKDFVLQNHNLNSNA 169  
 QY 61 DRIFDKTPPEP 70  
 DB 170 QKIFORPPDP 179

RESULT 13

VAL1\_TYLCV STANDARD: PRT; 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein (CI protein).  
 GN CI.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92024070; PubMed-1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RL with a single genomic component."  
 RL Virology 185:151-161(1991).

```

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X15656; CAA33688.1; -
CC PIR: D40779; OQCVCI.
CC InterPro: IPR001191; Gemini_AL1.
CC Pfam: PF00799; Gemini_AL1.1.
CC PRINTS: PR00227; GEMCOATALL.
CC ProDom: PD000736; Gemini_AL1.1.
CC ATP-binding.
CC NP_BIND 219 226 ATP (POTENTIAL).
CC SEQUENCE 357 AA; 40678 MW; 939AB68E1A83B2A7 CRC64;
SQ

```

Query Match 56.5%; Score 203; DB 1; Length 357;  
 Best Local Similarity 65.6%; Pred. No. 1e-15;  
 Matches 40; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

```

OY 4 MGSEFYDGRSARGCCQTSNDAAAEALNASSKEEALQIREKIPPAALFQFHNLNSNDRI 63
DB 111 FGVSDIGRSARGGGSANDAVAEALNAGSKSEALNLIKRAKAPDYTLQFHNLNSNDRI 170
OY 64 F 64
DB 171 F 171

```

RESULT 14

```

VALI_SICV STANDARD; PRT; 347 AA.
ID VALI_SICV
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE AL1 protein.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082449; PubMed=1984668;
RA Lazarowitz S.G.; Lazdins J.B.;
RT Infectivity and complete nucleotide sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype";
RL Virology 180:58-69(1991).

```

-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: M38183; AAC32410.1; ALT\_INT.  
 PIR: G36785; OQCVSI.  
 InterPro: IPR001191; Gemini\_AL1.  
 Pfam: PF00799; Gemini\_AL1.1.  
 PRINTS: PR00227; GEMCOATALL.  
 ProDom: PD000736; Gemini\_AL1.1.  
 ATP-binding.  
 NP\_BIND 218 225 ATP (POTENTIAL).  
 SEQUENCE 347 AA; 39110 MW; AFDABDEB122110E CRC64;

```

Query Match 32.9%; Score 118; DB 1; Length 347;
Best Local Similarity 37.9%; Pred. No. 4.2e-06;
Matches 25; Conservative 12; Mismatches 25; Indels 4; Gaps 1;
OY 5 GEQVDCRSARGCCQTSNDAAAEALNASSKEEALQIREKIPPAALFQFHNLNSNDRI 64
DB 116 GQYVSG-----GSKSNKDDYHNAAVNAAGSAGALDIKAGDKPTFIVYHNLLAVRLEF 171
OY 65 DKRPEP 70
DB 172 QKRPEP 177

```

RESULT 15

```

VEI_HPV66 STANDARD; PRT; 630 AA.
ID VEI_HPV66
AC Q80957;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Replication protein E1.
OS Human papillomavirus type 66.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=371119;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U31794; AAA79501.1; -
CC InterPro: IPR001177; Papillom_E1.
CC Pfam: PF00519; E1.1.
CC PRINTS: PR00524; E1_N; 1.
CC Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
CC Nuclear protein.
CC NP_BIND 458 465 ATP (POTENTIAL).
CC SEQUENCE 630 AA; 72065 MW; 22DDDA5934F7291B CRC64;
SQ

```

Query Match 18.1%; Score 65; DB 1; Length 630;  
 Best Local Similarity 25.5%; Pred. No. 8.3;  
 Matches 13; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

```

OY 19 QTSNDAAAEALNASSKEEALQIREKIPPAALFQFHNLNSNDRI 69
DB 269 KTRKSLSSILNVPDEOMLIQPKRLNSPAVALFYFTANSNSEVYGEPE 319

```

Search completed: October 10, 2002, 02:09:52  
 Job time: 1350 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:43 ; Search time 69.63 Seconds  
(without alignments)  
96.600 Million cell updates/sec

Title: US-09-289-346A-5  
Perfect score: 359  
Sequence: 1 TLVWGEFQVDCRSARSGCOT.....FOFHNLSNLDRIFFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR-71.\*

1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	343	99.5	352	1	QOCVLI
2	247	68.8	361	1	QOCVPT
3	243	67.7	358	1	S07594
4	232	64.6	362	1	J01887
5	226	63.0	359	2	S39211
6	224	62.4	349	2	J02300
7	224	62.4	349	2	S31875
8	219	61.0	359	2	S22593
9	216	60.2	351	1	J02327
10	216	60.2	358	1	J01870
11	215	59.9	355	1	QOCVMI
12	213	59.3	359	2	S39235
13	212	59.1	385	2	S28360
14	209	58.2	360	2	S59885
15	203	56.5	357	1	QOCVCI
16	134	37.3	131	1	S45059
17	118	32.9	347	1	QOCVSI
18	65	18.1	587	2	JC1419
19	63.5	17.7	1033	2	E97700
20	63.5	17.7	1713	2	A55347
21	62.5	17.4	642	2	D90558
22	62.5	17.3	160	2	G82060
23	62	17.3	308	2	E86840
24	61.5	17.1	300	2	G89582
25	61.5	17.1	538	2	AD0285
26	61	17.0	631	2	S36055
27	61	17.0	840	2	T36175
28	60.5	16.9	338	2	AG3582
29	60.5	16.9	429	2	A44384

## ALIGNMENTS

30	60.5	16.9	435	2	E59096	hypothetical prote
31	60.5	16.9	447	2	T12544	hypothetical prote
32	60.5	16.9	1070	1	A54600	1-phosphatidylinos
33	60	16.7	397	2	B71078	probable NADH oxid
34	60	16.7	714	2	C95382	probable ferrichro
35	59.5	16.6	706	2	H71707	hypothetical prote
36	59.5	16.6	1053	2	T06483	probable ubiquitin
37	59.5	16.6	1610	2	A46227	voltage-dependent
38	59.5	16.6	1646	2	JH0422	voltage-dependent
39	59.5	16.6	2137	2	T05244	hypothetical prote
40	59.5	16.6	2161	2	JH0564	calcium channel al
41	59.5	16.6	2181	2	A38198	calcium channel al
42	59.5	16.6	2203	2	T42742	voltage-dependent
43	59	16.4	351	2	E97408	fbpa (AE006039) [1
44	59	16.4	351	2	AF2626	hypothetical prote
45	59	16.4	451	2	S26839	retrovirus-related

## RESULT 1

QOCVLI

AL1 protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.5% Score 343; DB 1; Length 352;  
Best local Similarity 95.7% Pred. No. 1.6e-31;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDCRSARSGCOTSDNAAEALNMSKKEALQIIRKIPAAALFOFHNLSNL 60  
Db 111 TLVWGEFQVDCRSARSGCOTSDNAAEALNMSKKEALQIIRKIPKXYLFOFHNLSNL 170

QY 61 DRIFDKTPEP 70  
Db 171 DRIFDKTPEP 180

## RESULT 2

QOCVPT

AL1 protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: J00362; MUID:91311403

A:Accession: J00364

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-361 <CON>

A:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 68.8%; Score 247; DB 1; Length 361;  
 Best Local Similarity 68.1%; Pred. No. 1.6e-20;  
 Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRARGCGCOTSDNDAAEALNASSKEEALQIIRKIPAAALFOFHNLSNL 60  
 DB 110 TIEWGLFQIDGRSARGCGQTVNDAAAEALNASTGKEAMKTIKKELEKFLFOYHNLSNL 169  
 QY 61 DRIFDKTPE 69  
 DB 170 DRIFDKAPE 178

RESULT 3  
 507594  
 hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)  
 C:Species: cassava latent virus  
 C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
 C:Accession: S07594  
 R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.  
 Nucleic Acids Res. 18, 197-198, 1990  
 A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava  
 A:Reference number: S07590; MUID:90174930  
 A:Accession: S07594  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <MOR>  
 A:Cross-references: EMBL:X17095; NID:959371; PIDN:CAA34953.1; PID:959376  
 C:Genetics:  
 A:Map position: segment DNA1  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 67.7%; Score 243; DB 2; Length 358;  
 Best Local Similarity 64.3%; Pred. No. 4.5e-20;  
 Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRARGCGCOTSDNDAAEALNASSKEEALQIIRKIPAAALFOFHNLSNL 60  
 DB 109 TYEWGQFQIDGRSARGCGQSDANDAYAKALNASSKEALNVIRELPDVFLOFHNLSNL 168  
 QY 61 DRIFDKTPE 70  
 DB 169 DRIFDEPPAP 178

RESULT 4  
 J01887  
 AL1 protein - tomato yellow leaf curl virus (strain Australia)  
 N:Alternate names: CI protein  
 C:Species: tomato yellow leaf curl virus  
 C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01887  
 R:Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.  
 J. Gen. Virol. 74, 147-151, 1993  
 A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.  
 A:Reference number: J01885; MUID:9313978  
 A:Accession: J01887  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-362 <DRY>  
 A:Cross-references: GB:S53251  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 64.6%; Score 232; DB 1; Length 362;  
 Best Local Similarity 55.3%; Pred. No. 8.2e-19;  
 Matches 47; Conservative 8; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLWGEFQVDSRARGCGCOTSDNDAAEALNASSKEEALQIIRKIPAAALFOFHNLSNL 60  
 DB 110 TIEWGLFQIDGRSARGCGQTVNDAAAEALNASTGKEAMKTIKKELEKFLFOYHNLSNL 169

DB 110 TIEWGEFQIDGRSARGCGQSDANDAYAOALNTGSKSEALNVIRELPAPKDYVLOFHNLSNL 169  
 QY 61 DRI-----FDKTPPE 69  
 DB 170 DRIFPLEVYVSPELSSSFDRAPE 194

RESULT 5  
 S39211  
 gene CI protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
 C:Accession: S39211  
 R:Morris, E.; Hida190, E.; Accotto, G.; Moriones, E.  
 submitted to the EMBL Data Library, August 1993  
 A:Description: High similarity among the tomato yellow leaf curl virus isolates from  
 A:Reference number: S39209  
 A:Accession: S39211  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <NOR>  
 A:Cross-references: EMBL:Z25751; NID:9433655; PIDN:CAA81026.1; PID:9433658  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 63.0%; Score 226; DB 2; Length 359;  
 Best Local Similarity 60.9%; Pred. No. 3.9e-18;  
 Matches 42; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 2 TLWGEFQVDSRARGCGCOTSDNDAAEALNASSKEEALQIIRKIPAAALFOFHNLSNL 61  
 DB 111 LEMGTQFQIDGRSARGCGQTVNDAYAKAINAGSKSEALDVIRELPDVFLOFHNLSNL 170  
 QY 62 DRIFDKTPE 70  
 DB 171 RVFQVPPAP 179

RESULT 6  
 J02300  
 replicase - pepper huasteco virus (component A)  
 N:Alternate names: ORF AL1 protein  
 C:Species: pepper huasteco virus  
 C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999  
 C:Accession: J02300  
 R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,  
 J. Gen. Virol. 74, 2225-2231, 1993  
 A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b  
 A:Reference number: J02299; MUID:94015007  
 A:Accession: J02300  
 A:Molecule type: DNA  
 A:Residues: 1-349 <TOR>  
 A:Cross-references: GB:X70418; NID:961023; PIDN:CAA49856.1; PID:961025  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 62.4%; Score 224; DB 2; Length 349;  
 Best Local Similarity 61.4%; Pred. No. 6.5e-18;  
 Matches 43; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRARGCGCOTSDNDAAEALNASSKEEALQIIRKIPAAALFOFHNLSNL 60  
 DB 110 TIEWGEFQIDGRSARGCGQSDANDAYAKALNASSKEEALQIIRKIPAAALFOFHNLSNL 169  
 QY 61 DRIFDKTPE 70  
 DB 170 NRIFOTPPPE 179

RESULT 7  
 S31875  
 AL1 protein - pepper rizado amarillo virus  
 C:Species: pepper rizado amarillo virus

A; Accession: JQ2327

Best Local Similarity 58.68; Pred. No. 7.1e-17;

Matches	41;	Conservative	11;	Mismatches	18;	Indels	0;	Gaps	0;
---------	-----	--------------	-----	------------	-----	--------	----	------	----

```

OY      1  TLWGEQVQVGRSARGGCCQTSNIAAEALVASSKEKEALIIIREKIIPAALFOFHNIENL  60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      110 TAEWGEFQIDGRSARSGQQTANDSYAKALNAGDVQVQALNLTKEQGRDQVVLQNHNIENL  169
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      61 DRFDKTPR 70
        : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      170 ERIFAKAPER 179
        : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 12  
S39235  
gene C1 protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
C:Accession: S39235  
R: Crespi, S.; Morris, E.; Vaira, A.; Bosco, D.; Accotto, G.  
submitted to the EMBL Data Library, December 1993  
A: Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity  
A: Reference number: S39233  
A: Accession: S39235  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-359 <CRE>  
A: Cross-references: EMBL: Z28390; NID: q1041671; PTD: q1334964  
C: Superfamily: tomato golden mosaic virus AL1 protein

Query Match	59.3%;	Score 213;	DB 2;	length 359;
Best Local Similarity	56.5%;	Pred. No. 1.2e+16;		
Matches 39;	Conservative 12;	Mismatches 18;	Indels 0;	Gaps 0;

[illegible]

RESULT 13  
S28360  
AL1 protein - beet curly top virus  
C:Species: beet curly top virus  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
C:Accession: S28360  
R:Stanley J., Markham P.G., Callis, R.J., Pinner, M.S.  
EMBO J. 5, 1761-1767, 1986  
A:title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top  
A:Reference number: S28360  
A:Accession: S28360  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-385 <STA>  
A:Cross-references: GB:M24597; EMBL:X04144; NID:9210678; PIDN:AAA42751.1; PID:9210679  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	59.1%	Score 212	DB 2:	Length 385
Best Local Similarity	55.7%	Pred. No. 1.7e-16		
Matches 39; Conservative	15;	Mismatches 16;	Indels 0;	Gaps 0

[illegible]

```

559885      replication-associated protein C1 - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C>Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C:Accession: S59885
R:Hong, Y.; Harrison, B.D.
Submitted to the EMBL Data Library, February 1995
d.Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d.geminiViruses.
A:Reference number: S58346
A:Accession: S59885
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1..360 <HON>
A:Cross-references: EMBL:Z48182; NID:G944838; PIDN:CAA88229.1; PID:9974211
C:superfamily: tomato golden mosaic virus At1 protein

Query Match          58.2%, Score 209, DB 2, Length 360;
Best Local Similarity 62.1%; Pred. No. 3.5e-16;
Matches    41; Conservative   10; Mismatches   15; Indels       0; Gaps     0;

QY      4 WGEFYVDGRSARGGCOTSNDDAAAEALNNSSKEALQITREKIPPAALFOFHTLNENLDRI 63
      :|||:|||:|||::|||:|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      113 FGVFPIDGSGARGGQSANDAYAEAINSGSKAAALDILREKAPKFVLOFHTLNANLDRI 172
      :|||:|||:|||::|||:|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      64 FDKTPE 69
      | : |
Db      173 ETTPSAE 178

```

Query Match	58.28;	Score 209;	DB 2;	Length 360;
Best Local Similarity	62.18;	Pred. No. 3.5e-16;		
Matches 41; Conservative	10;	Mismatches 15;	Indels 0;	Gaps 0;

OY	4	WEEFVDDRSRGCGQTSDNDAALNNSSKEEALQIIRKIPAAAFQCHLNLSNDRI	63
	:   :	:   :	:   :
Db	113	FGEVFIDGRSARGGQSANDAYAAAINSGSKAALDILREKAPDFVLOFHNLNANLDRI	172
OY	64	FDKTPTE	69
	:   :		
Db	173	FTPSAE	178

RESULT 15  
OOCVC1  
A1: protein - tomato yellow leaf curl virus  
N: Alternate names: CI protein  
C: Species: tomato yellow leaf curl virus  
C: Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
C: Accession: D40779  
R: Navot, N.; Pichersky, E.; Zeldan, M.; Zamir, D.; Czosnek, H.  
Virology 185, 151-161, 1991  
A: Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin-  
A: Accession: D40779; MUID: 92024070  
A: Reference: translation not shown  
A: Molecule type: DNA  
A: Residues: 1-357 <NAV>  
A: Cross-references: GB:X15656; NID:962204; PIDN:CA3688.1; PID:962207  
C: Superfamily: tomato golden mosaic virus A1 protein

Query Match	56.5%	Score 203	DB 1	Length 357
Similarity	65.6%	Pred. No.	1.7e-15	
Best Local	68			
Matches	40	Mismatches	12	Indels 0
				Gaps 0

QY 4 WGEFQVQDGRSARGGCQTSNDAAAEALMASSKEEAQIIREKIPAAALFQFHNLNSNDRI 63  
: | : | | | | | | | : | | | | | | | : | | | | | | | : | | | | | | | |  
Db 111 FGVSQIDGRSARGGQGSANDAYAEALNSSKSEALNIKEAPKDYILQFHNLSNDRI 170

QY	64	F	64
		+	
Db	171	F	171

Search completed: October 10, 2002, 01:48:44  
Job time: 562 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:02 ; Search time 407.82 Seconds  
(Without alignments)  
47.487 Million cell updates/sec

Title: US-09-289-346A-5

Perfect score: 359

Sequence: 1 TLVWGFQVDCRSARCGCOT.....PQPHNLSNLDRIPTTPP 70

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1006125 seqs, 27659714 residues

Total number of hits satisfying chosen parameters: 1006125

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCR\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	19.8	340	6	US-10-219-999-59588
2	66.5	18.5	512	5	US-09-791-537-52432
3	66.5	18.5	1449	5	US-09-791-537-52442
4	65	18.1	587	5	US-09-791-537-33198
5	63.5	17.7	1238	5	US-09-791-537-14585
6	63.5	17.7	1331	5	US-09-791-537-118290
7	63.5	17.7	1713	5	US-09-643-992A-2
8	63.5	17.7	1713	5	US-09-791-537-87556
9	63.5	17.7	1806	5	US-10-171-311-113
10	63.5	17.7	1806	5	US-09-791-537-53064
11	63.5	17.7	3298	7	US-60-365-384-181
12	63.5	17.7	3332	7	US-60-365-384-182
13	62.5	17.4	180	6	US-10-155-881-31455
14	62.5	17.4	199	6	US-10-155-881-18282
15	62.5	17.4	429	5	US-09-791-537-30006
16	62.5	17.4	434	6	US-10-179-131-5314
17	62.5	17.4	945	5	US-09-791-537-132140
18	62.5	17.4	300	5	US-09-791-537-98310
19	61.5	17.1	300	7	US-60-360-039-6839
20	61.5	17.1	365	6	US-10-155-881-28183
21	61.5	17.1	365	7	US-60-411-837-483
22	61.5	17.1	475	7	US-60-389-987-1438
23	61.5	17.1	476	5	US-09-582-687A-62
24	60.5	16.9	429	5	US-09-791-537-53935
25	60.5	16.9	435	5	US-09-791-537-147074
26	60.5	16.9	447	5	US-09-991-150-369

27	60.5	16.9	447	5	US-09-989-328-369	Sequence 369, App
28	60.5	16.9	447	6	US-10-103-295-129	Sequence 129, App
29	60.5	16.9	447	6	US-10-121-062-254	Sequence 254, App
30	60.5	16.9	447	6	US-10-173-689-254	Sequence 254, App
31	60.5	16.9	447	6	US-10-173-690-254	Sequence 254, App
32	60.5	16.9	447	6	US-10-173-691-254	Sequence 254, App
33	60.5	16.9	447	6	US-10-173-692-254	Sequence 254, App
34	60.5	16.9	447	6	US-10-173-695-254	Sequence 254, App
35	60.5	16.9	447	6	US-10-173-697-254	Sequence 254, App
36	60.5	16.9	447	6	US-10-173-698-254	Sequence 254, App
37	60.5	16.9	447	6	US-10-173-699-254	Sequence 254, App
38	60.5	16.9	447	6	US-10-173-701-254	Sequence 254, App
39	60.5	16.9	447	6	US-10-173-704-254	Sequence 254, App
40	60.5	16.9	447	6	US-10-173-708-254	Sequence 254, App
41	60.5	16.9	447	6	US-10-174-571-254	Sequence 254, App
42	60.5	16.9	447	6	US-10-174-574-254	Sequence 254, App
43	60.5	16.9	447	6	US-10-174-583-254	Sequence 254, App
44	60.5	16.9	447	6	US-10-173-694-254	Sequence 254, App
45	60.5	16.9	447	6	US-10-173-707-254	Sequence 254, App

## ALIGNMENTS

RESULT 1  
: Sequence 59588, Application US/10219999  
: GENERAL INFORMATION:  
: APPLICANT: Cao, Yongwei  
: APPLICANT: Edgerton, Michael D  
: APPLICANT: Hinkle, Gregory J.  
: APPLICANT: Kovacic, David K.  
: APPLICANT: Liu, Jingdong  
: APPLICANT: Stein, Joshua  
: TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
: FILE REFERENCE: 38-10(52726)C  
: CURRENT APPLICATION NUMBER: US/10/219, 999  
: PRIOR FILING DATE: 2002-08-15  
: PRIOR APPLICATION NUMBER: US 60/324,109  
: PRIOR FILING DATE: 2001-09-21  
: PRIOR APPLICATION NUMBER: US 60/312,544  
: PRIOR FILING DATE: 2001-08-15  
: NUMBER OF SEQ ID NOS: 63520  
: SEQ ID NO 59588  
: LENGTH: 340  
: TYPE: PRT  
: ORGANISM: Zea mays  
US-10-219-999-59588  
Query Match 19.8%; Score 71; DB 6; Length 340;  
Best local similarity 28.2%; Pred. No. 1.9;  
Matches 20; Conservative 13; Mismatches 22; Indels 16; Gaps 3;  
OY 15 RGGCTSDAAAEAL-NSSKEBALQ-----LIREKIPAAALFOFHN-----LNS 58  
DB 30 RFGCSSRPRPTATVNHGATKQVROCKVSDETKMWDELPQAWMFDSHNLARSPWLNS 89  
OY 59 NIDRFDKTPE 69  
DB 90 TLSELDKTKQ 100  
RESULT 2  
US-09-791-537-52432  
: Sequence 52432, Application US/09791537  
: GENERAL INFORMATION:  
: APPLICANT: Biomimix, Inc.  
: APPLICANT: Depe, Derek  
: APPLICANT: Dazzer, Joseph  
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
: FILE REFERENCE: 261/210

Query Match	17.7%;	Score 63.5;	DB 5;	Length 1331;
Best Local Similarity	34.4%;	Pred. No. 95;		
Matches	21; Conservative	12; Mismatches	21; Indels	7; Gaps 3



QY 9 VDGSRAGGCGTSDNDAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSLDRIFD 65  
 DB 646 VDAATAYENILNIAKAEADANRAASASESALQTVIKEDLPRA-----KTLSSNSDKLNL 701  
 QY 66 K 66  
 DB 702 E 702

RESULT 7  
 US-09-643-992A-2  
 ; Sequence 2, Application US/09643992A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Oriental Yeast Co., Ltd.  
 ; APPLICANT: MATUO Toshikuni  
 ; TITLE OF INVENTION: Method for neuragenesis  
 ; FILE REFERENCE: EP/0-21-8US  
 ; CURRENT APPLICATION NUMBER: US/09/643,992A  
 ; CURRENT FILING DATE: 2000-08-23  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SEQ ID NO 2  
 ; LENGTH: 1713  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-643-992A-2

Query Match 17.7%; Score 63.5; DB 5; Length 1713;  
 Best Local Similarity 34.4%; Pred. No. 1.3e+02;  
 Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

QY 9 VDGSRAGGCGTSDNDAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSLDRIFD 65  
 DB 553 VDAATAYENILNIAKAEADANRAASASESALQTVIKEDLPRA-----KTLSSNSDKLNL 608  
 QY 66 K 66  
 DB 609 E 609

RESULT 8  
 US-09-791-537-87556  
 ; Sequence 87556, Application US/09791537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bionomix, Inc.  
 ; APPLICANT: Debe, Derek  
 ; APPLICANT: Danzer, Joseph  
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
 ; TITLE OF INVENTION: METHODS OF USE THEREOF  
 ; FILE REFERENCE: 261/210  
 ; CURRENT APPLICATION NUMBER: US/09/791,537  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 153055  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 87556  
 ; LENGTH: 1713  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-791-537-87556

Query Match 17.7%; Score 63.5; DB 5; Length 1713;  
 Best Local Similarity 34.4%; Pred. No. 1.3e+02;  
 Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

QY 9 VDGSRAGGCGTSDNDAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSLDRIFD 65  
 DB 553 VDAATAYENILNIAKAEADANRAASASESALQTVIKEDLPRA-----KTLSSNSDKLNL 608  
 QY 66 K 66  
 DB 609 E 609

RESULT 9  
 US-10-171-311-113  
 ; Sequence 113, Application US/10171311  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schlegel, Robert  
 ; APPLICANT: Chen, Yan  
 ; APPLICANT: Zhao, Xumei  
 ; APPLICANT: Monahan, John  
 ; APPLICANT: Kamalkar, Shubhangi  
 ; APPLICANT: Glatz, Karen  
 ; APPLICANT: Gannavarapu, Manjula  
 ; APPLICANT: Hoersch, Sebastian  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
 ; FILE REFERENCE: MRI-035  
 ; CURRENT APPLICATION NUMBER: US/10/171,311  
 ; CURRENT FILING DATE: 2002-06-12  
 ; PRIOR APPLICATION NUMBER: US 60/298,159  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,155  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/335,936  
 ; PRIOR FILING DATE: 2001-11-14  
 ; NUMBER OF SEQ ID NOS: 238  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 113  
 ; LENGTH: 1713  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-171-311-113

Query Match 17.7%; Score 63.5; DB 6; Length 1713;  
 Best Local Similarity 34.4%; Pred. No. 1.3e+02;  
 Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

QY 9 VDGSRAGGCGTSDNDAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSLDRIFD 65  
 DB 553 VDAATAYENILNIAKAEADANRAASASESALQTVIKEDLPRA-----KTLSSNSDKLNL 608  
 QY 66 K 66  
 DB 609 E 609

RESULT 10  
 US-09-791-537-53064  
 ; Sequence 53064, Application US/09791537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bionomix, Inc.  
 ; APPLICANT: Debe, Derek  
 ; APPLICANT: Danzer, Joseph  
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
 ; TITLE OF INVENTION: METHODS OF USE THEREOF  
 ; FILE REFERENCE: 261/210  
 ; CURRENT APPLICATION NUMBER: US/09/791,537  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 153055  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 53064  
 ; LENGTH: 1806  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-791-537-53064

Query Match 17.7%; Score 63.5; DB 5; Length 1806;  
 Best Local Similarity 34.4%; Pred. No. 1.4e+02;  
 Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```

RESULT 14
US-10-155-881-18282
; Sequence 18282, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovacic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NOCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

```

```

: TITLE OF INVENTION: TRANSCRIPTION IN PLANTS
: FILE REFERENCE: 38-21(15300)J
: CURRENT APPLICATION NUMBER: US/10/155,881
: CURRENT FILING DATE: 2002-05-22
: NUMBER OF SEQ ID NOS: 37595
: SEQ ID NO 18282
: LENGTH: 199
: TYPE: PRT
: ORGANISM: Glycine max
US-10-155-881-18282

```

```

Query Match          17.4%  Score 62.5; DB 6; Length 199;
Best Local Similarity 34.3%  Pred. No. 12;
Matches 23; Conservative 9; Mismatches 24; Indels 11; Gaps 3;

```

```

QY      4  WGEFQVDGR-SARGG-----QTSNDAAAEALNASSK---EEALQIIREKIPAAALFQ 52
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      99  WCKYAAEIRDSARHGARIWLTGTFQTAEEAAMAYDRAAFKMGSKALLNFPALIVATELS 158
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      53  FHNLSN 59
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      159  FNNTNSN 165
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 15
US-09-791-537-30006
: Sequence 30006, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 30006
: LENGTH: 429
: TYPE: PRT
: ORGANISM: Candida albicans
US-09-791-537-30006

```

```

Query Match          17.4%  Score 62.5; DB 5; Length 429;
Best Local Similarity 29.0%  Pred. No. 31;
Matches 20; Conservative 16; Mismatches 24; Indels 9; Gaps 3;

```

```

QY      1  TLVWGEFQVDGRSARGGCGTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 60
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      162  TDIWGR---DDDS-----NNSDAINQALESSLNKDSQFTRLSI-AEAIHKLWKLDGSI 212
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      61  DRIFDKTPE 69
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      213  KKCPRDSNE 221
      111: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: October 10, 2002, 02:09:02  
Job time: 1445 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 : Search time 133.15 Seconds  
(without alignments)  
58.394 Million cell updates/sec

Title: US-09-289-346a-5

Perfect score: 359

Sequence: 1 TLVWGEFQVDGSRAGCCOT.....PQHNLSNLDRIKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	70	21	AA18681
2	343	95.5	70	21	AA18677
3	343	95.5	70	21	AA18687
4	335	93.3	70	21	AA18685
5	332	92.5	70	21	AA18688
6	331	92.2	70	21	AA18692
7	329	91.6	70	21	AA18684
8	329	91.6	70	21	AA18690
9	328	91.4	70	21	AA18678
10	328	91.4	70	21	AA18686
11	327	91.1	70	21	AA18689

12	325	90.5	70	21	AA18680	Mutant peptide der
13	325	90.5	70	21	AA18691	Mutant peptide der
14	321	89.4	70	21	AA18683	Mutant peptide der
15	319	88.9	70	21	AA18682	Mutant peptide der
16	311	86.6	70	21	AA18679	Mutant peptide der
17	220	61.3	361	18	AAW34336	Tomato mottle viru
18	220	61.3	361	18	AAW34324	Tomato mottle viru
19	220	61.3	361	18	AAW34325	Tomato mottle viru
20	220	61.3	361	18	AAW34326	Tomato mottle viru
21	219	61.0	359	17	AAW34327	Sardinian tomato y
22	219	61.0	359	17	AAW34328	Sardinian tomato y
23	219	61.0	359	17	AAW34329	Bean golden mosaic
24	218	60.7	353	18	AAW34330	Bean golden mosaic
25	218	60.7	353	18	AAW34331	Bean golden mosaic
26	218	60.7	353	18	AAW34332	Bean golden mosaic
27	218	60.7	353	18	AAW34333	Bean golden mosaic
28	218	60.7	353	18	AAW34334	Bean golden mosaic
29	216	60.2	353	18	AAW34335	ORF 4 gene product
30	208.5	58.1	361	8	AAW70562	Product of ORF 4 f
31	208	57.9	362	19	AAW56495	Tomato yellow leaf
32	203	56.5	357	18	AAW34329	Tomato yellow leaf
33	203	56.5	357	18	AAW34330	Tomato yellow leaf
34	203	56.5	357	18	AAW34331	Tomato yellow leaf
35	195	54.3	357	18	AAW34337	Tomato yellow leaf
36	66.5	18.5	512	19	AAW68473	HTV-1 strain YBF30
37	65.5	18.2	665	22	ABB70991	Drosophila melanog
38	63.5	17.7	1693	21	AAW48457	Human laminin 5 po
39	63.5	17.7	1693	21	AAW48459	Human laminin 5 po
40	63.5	17.7	1713	16	AAW70148	Deduced sequence o
41	63.5	17.7	1713	16	AAW48458	Human laminin 5 po
42	63.5	17.7	1724	21	AAW48456	Novel human diagno
43	63	17.5	599	22	ABG14172	Novel human secret
44	63	17.5	599	22	AAU32811	Drosophila melanog
45	62.5	17.4	945	22	ABB65231	

## ALIGNMENTS

RESULT 1	AA18681	standard; peptide; 70 AA.
ID	AA18681	
XX	AA18681	
AC	AA18681	
XX	22-JAN-2001	(first entry)
DT		
XX		
DE		Mutant peptide derived from amino acids 110-179 of Rep (All) protein.
XX		
DE		Geminivirus; replication protein; Rep protein; All; transgenic plant;
XX		
KW		ribosome binding region; resistance; geminivirus infection.
XX		
OS		Synthetic.
XX		
OS		Tomato golden mosaic virus.
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 47	
FT	Misc-difference 48	/note= "wild type residue replaced with Ala"
FT	Misc-difference 49	/note= "wild type residue replaced with Ala"
FT	Misc-difference 49	/note= "wild type residue replaced with Ala"
XX		
PN	WO200054573-A1.	
XX		
PD	21-SEP-2000.	
XX		
PF	15-MAR-2000; 2000WO-0506759.	
XX		
PR	18-MAR-1999; 99US-0125004.	
XX		
PR	09-APR-1999; 99US-0289346.	
XX		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
PI WPI; 2000-618851/59.

PT Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
PS Claim 52; Page 44; 73pp; English.

XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.

XX Sequence 70 AA;

Query Match 100.0%; Score 359; DB 21; Length 70;  
Best Local Similarity 100.0%; Pred. No. 3.5e-39;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGRSARGCGQTSDNAAAEALNASSKEEALQITREKIPPAALFOFNLSNL 60  
DB 1 tlwgeftqvdgrsargcgqtsndaaaealnasskeeaqlitrekippaalfqfnlnsnl 60

QY 61 DRFDKTPPP 70  
DB 61 drfdktppp 70

#### RESULT 2

ID AAB18677 standard; peptide; 70 AA.

XX AAB18677;

DT 22-JAN-2001 (first entry)

DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).

KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

KW ribosome binding region; resistance; geminivirus infection.

OS Tomato golden mosaic virus.

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

PI Hanley-Bowdoin L, Orozco BM, Kong L;

DR WPI; 2000-618851/59.

XX

PT Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
PS Disclosure; Page 18; 73pp; English.

XX The present sequence is derived from a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.

XX Sequence 70 AA;

Query Match 95.5%; Score 343; DB 21; Length 70;  
Best Local Similarity 95.7%; Pred. No. 4.3e-37;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGRSARGCGQTSDNAAAEALNASSKEEALQITREKIPPAALFOFNLSNL 60  
DB 1 tlwgeftqvdgrsargcgqtsndaaaealnasskeeaqlitrekippaalfqfnlnsnl 60

QY 61 DRFDKTPPP 70  
DB 61 drfdktppp 70

#### RESULT 3

ID AAB18687 standard; peptide; 356 AA.

XX AAB18687;

DT 22-JAN-2001 (first entry)

DE Amino acid sequence of a geminivirus replication protein of TGMV.

KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

KW ribosome binding region; resistance; geminivirus infection.

OS Tomato golden mosaic virus.

FN Key location/Qualifiers

FT Misc-difference 354 /note= "unspecified amino acid"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

PI Hanley-Bowdoin L, Orozco BM, Kong L;

DR WPI; 2000-618851/59.

PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PS  
 XX  
 PS Disclosure; Page 47-48; 73pp; English.

CC The present sequence represents a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX  
 SQ Sequence 356 AA;

Query Match 95.5%; Score 343; DB 21; Length 356;  
 Best Local Similarity 95.7%; Pred. No. 3,7e-36;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGCGCOTSDNDAAEALNASKKEALQIREKIPAAALFQPHNLSNL 60  
 |||||  
 DB 110 TLVGEFQVDGRSARGCGCOTSDNDAAEALNASKKEALQIREKIPAAALFQPHNLSNL 169

QY 61 DRIFDKTPEP 70  
 |||||  
 DB 170 DRIFDKTPEP 179

RESULT 4  
 AAB18685  
 ID AAB18685 standard; peptide: 70 AA.  
 AC AAB18685;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 XX Synthetic.  
 OS Tomato golden mosaic virus.

XX  
 FH Key Location/Qualifiers  
 FT MISC-difference 10  
 FT MISC-difference 20 /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI  
 XX WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT

PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PS  
 XX  
 PS Claim 53; Page 46; 73pp; English.

CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;

Query Match 93.3%; Score 335; DB 21; Length 70;  
 Best Local Similarity 94.3%; Pred. No. 4.8e-36;  
 Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGCGCOTSDNDAAEALNASKKEALQIREKIPAAALFQPHNLSNL 60  
 |||||  
 DB 1 TLVGEFQVDGRSARGCGCOTSDNDAAEALNASKKEALQIREKIPAAALFQPHNLSNL 60

QY 61 DRIFDKTPEP 70  
 |||||  
 DB 61 DRIFDKTPEP 70

RESULT 5  
 AAB18688  
 ID AAB18688 standard; peptide: 70 AA.  
 AC AAB18688;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 XX Synthetic.  
 OS Tomato golden mosaic virus.

XX  
 FH Key Location/Qualifiers  
 FT MISC-difference 19  
 FT MISC-difference 20 /note= "wild type residue replaced with Ala"  
 FT MISC-difference 20 /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI  
 XX WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT

XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
PS Disclosure; Page 48; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA:

Query Match 92.5%; Score 332; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 1.2e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFYVDGRSARGCQTSDNDAAEALNASSKEEALQIREKIPPAALFOFHNLNSL 60  
Db 1 tlwvgefyvdgrsargcqtndaaealnasckealqirekipekyifqfhnlnsl 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

RESULT 6  
AAB18692  
ID AAB18692 standard; peptide; 70 AA.  
XX  
AC AAB18692;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

OS Tomato golden mosaic virus.

XX Key Location/Qualifiers

FT Misc-difference 66 /note= "wild type residue replaced with Ala"

FT Misc-difference 69 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

PI Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
PS Disclosure; Page 50; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA:

Query Match 92.2%; Score 331; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 1.6e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFYVDGRSARGCQTSDNDAAEALNASSKEEALQIREKIPPAALFOFHNLNSL 60  
Db 1 tlwvgefyvdgrsargcqtndaaealnasckealqirekipekyifqfhnlnsl 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

RESULT 7  
AAB18684  
ID AAB18684 standard; peptide; 70 AA.  
XX  
AC AAB18684;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

OS Tomato golden mosaic virus.

XX Key Location/Qualifiers

FT Misc-difference 7 /note= "wild type residue replaced with Ala"

FT Misc-difference 8 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.



XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT -  
 PS Claim 52: Page 45; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX Sequence 70 AA;

Query Match 91.6%; Score 329; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 2.9e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLWGEFOYDGRSARGCGQTSNDAAEALNASSKEEALQIREKIPPAALFQFHNLNSL 60  
 Db 1 tlwgeaavdgrrsargcgqtsndaaealnaskeaalqirekipekylfqfhnlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdkrpep 70

RESULT 8  
 AAB18690  
 ID AAB18690 standard; peptide: 70 AA.

AC AAB18690;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

KM ribosome binding region; resistance; geminivirus infection.

OS Synthetic.

OS Tomato golden mosaic virus.

FT Key Location/Qualifiers

FT Misc-difference 27 /note= "wild type residue replaced with Ala"

FT Misc-difference 30 /note= "wild type residue replaced with Ala"

FT Misc-difference 15 /note= "wild type residue replaced with Ala"

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UNNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence

PT encoding a mutant AL1 protein with a mutation in the Rb binding region

PT -  
 PS Disclosure; Page 49; 73pp; English.

XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX Sequence 70 AA;

Query Match 91.6%; Score 329; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 2.9e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLWGEFOYDGRSARGCGQTSNDAAEALNASSKEEALQIREKIPPAALFQFHNLNSL 60  
 Db 1 tlwgefgyvdgrsargcgqtsndaaealnaskeaalqirekipekylfqfhnlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdkrpep 70

RESULT 9

AAB18678  
 ID AAB18678 standard; peptide: 70 AA.

AC AAB18678;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

KM ribosome binding region; resistance; geminivirus infection.

OS Synthetic.

OS Tomato golden mosaic virus.

FT Key Location/Qualifiers

FT Misc-difference 12 /note= "wild type residue replaced with Ala"

FT Misc-difference 13 /note= "wild type residue replaced with Ala"

FT Misc-difference 15 /note= "wild type residue replaced with Ala"

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

```

PF 15-MAR-2000; 2000OWO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
PI Hanley-Bowdoin L, Orozco BM, Kong L;
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rd binding region
PT -
XX
PS Claim 53; Page 42-43; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 91.4%; Score 328; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 3.9e-35;
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0.

QY 1 TLVWGEFOYVGRRARGCQCTNSDDAAEAALMASKEEALQIRKTIIPAALFOPHNINSNL 60
   |||||
Db 1 tlvwgefvydgaaagcgctnsdaaeaalnasskeaalqirkiptkylfqtfnunsl 60
   |||||
QY 61 DRIFDKTPEP 70
   |||||||
Db 61 drifdktppep 70

RESULT 10
AAB18686
ID AAB18686 standard; peptide; 70 AA.
AC
XX
AC AAB18686;
XX
DT 22-JAN-2001 (first entry)
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
OS Synthetic.
XX Tomato golden mosaic virus.
FH Key Location/Qualifiers
FT Misc-difference 24 /note= "wild type residue replaced with leu"
FT Misc-difference 25 /note= "wild type residue replaced with leu"
FT FT /note= "wild type residue replaced with leu"
FT Misc-difference 26 /note= "wild type residue replaced with leu"
XX
```

[illegible]

FT /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000MO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 XX 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 DR  
 XX  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure; Page 48-49; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SO Sequence 70 AA;  
 Query Match 91.1%; Score 327; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 5.3e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPPAALFQFHNLNSL 60  
 DB 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPPAALFQFHNLNSL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 drifdktp 70  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 XX Tomato golden mosaic virus.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 42

FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 43  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 44  
 FT /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000MO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 XX 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 DR  
 XX  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 52; Page 43-44; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SO Sequence 70 AA;  
 Query Match 90.5%; Score 325; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 9.6e-35;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPPAALFQFHNLNSL 60  
 DB 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPPAALFQFHNLNSL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 drifdktp 70  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 42

OS	Tomato golden mosaic virus.
XX	
PH	Key Location/Qualifiers
FT	Misc-difference 34 /note= "wild type residue replaced with Ala"
FT	Misc-difference 35 /note= "wild type residue replaced with Ala"
FT	Misc-difference 36 /note= "wild type residue replaced with Ala"
XX	
PN	WO200054573-A1.
PD	
PD	21-SEP-2000.
PF	15-MAR-2000; 2000MC-US06759.
PR	18-MAR-1999; 99US-0125004.
PR	09-APR-1999; 99US-0289346.
PA	(UYNC-) UNIV NORTH CAROLINA STATE.
XX	
PI	Hanley-Bowdoin L, Orozco BM, Kong L;
XX	WPI; 2000-61851/59.
XX	
PT	Transgenic plants with increased resistance to geminivirus infection comprise a nucleic acid construct containing a nucleic acid sequence encoding a mutant AL1 protein with a mutation in the Rb binding region
PT	-
PS	Disclosure; Page 49; 73pp; English.
XX	
CC	The present sequence represents a mutant peptide, derived from a geminivirus replication (Rep) protein, also known as AL1. AL1 binds double-stranded DNA, catalyzes cleavage and ligation of single-stranded DNA, and interacts with other viral and host proteins. Mutants of the AL1 protein are used to produce transgenic plants. The mutation in AL1 is present in a ribosome binding region, and expression of mutant AL1 protein imparts increased resistance to geminivirus infection in the plant. Mutant AL1 proteins are useful for producing plants having increased resistance or reduced sensitivity to a geminivirus such as tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl virus, African cassava mosaic virus, Indian cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or beet curly top virus.
XX	
SQ	Sequence 70 AA:
	Query Match 90.5%; Score 325; DB 21; Length 70;
	Best Local Similarity 91.4%; Pred. No. 9, 6e-35;
	Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0
OY	1 TLVGEFOVDGRSARGCCTSDNDAAEALINASSKEEALOTIRKIPAAALFOPHNLSNI 60   Db 1 TLVGEFGVDGSRARGCQTNDAAEAALNASSAAAAIQLIREKYLIGTHMINSNI 60
OY	61 DRIEDKTPEP 70   Db 61 dritfdktpep 70
RESULT 14	
ID AAB18683	
XX AAB18683 standard; peptide; 70 AA.	
AC AAB18683;	
XX	
DT 22-JAN-2001 (first entry)	
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.	
XX	

KW	Geminivirus: replication protein; Rep protein; AL1: transgenic plant;			
XV	ribosome binding region; resistance; geminivirus infection.			
XX	Synthetic.			
OS	Tomato golden mosaic virus.			
XX				
FH	Key	Location/Qualifiers		
FT	Misc-difference	59	/note=	"wild type residue replaced with Ala"
FT	Misc-difference	61	/note=	"wild type residue replaced with Ala"
FT	Misc-difference	62	/note=	"wild type residue replaced with Ala"
PN	WO200054573-A1.			
XX	21-SEP-2000.			
XX	15-MAR-2000; 2000MO-US06759.			
XX	18-MAR-1999; 99US-0125004.			
PR	09-APR-1999; 99US-0289346.			
XX	(UYNC-) UNIV NORTH CAROLINA STATE.			
PA	Hanley-Bowdoin L, Orozco BW, Kong L;			
PI	WPI: 2000-618851/59.			
DR				
XX	Transgenic plants with increased resistance to geminivirus infection			
FT	comprise a nucleic acid construct containing a nucleic acid sequence			
PT	encoding a mutant AL1 protein with a mutation in the Rb binding region			
PT	-			
XX				
PS	Claim 53; Page 45; 73pp; English.			
XX				
CC	The present sequence represents a mutant peptide, derived from a			
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds			
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded			
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1			
CC	protein are used to produce transgenic plants. The mutation in AL1 is			
CC	present in a ribosome binding region, and expression of mutant AL1			
CC	protein imparts increased resistance to geminivirus infection in the			
CC	plant. Mutant AL1 proteins are useful for producing plants having			
CC	increased resistance or reduced sensitivity to a geminivirus such as			
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl			
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian			
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic			
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper			
CC	virus, cotton leaf curl virus or beet curly top virus.			
XX				
XX	Sequence	70 AA;		
XX				
XX	Query Match	89.4%;	Score 321;	DB 21; Length 70;
XX	Best Local Similarity	91.4%;	Pred. No. 3.2e-34;	
XX	Matches 64; Conservative	0;	Mismatches 6;	Indels 0; Gaps 0;
QY	1	TLVGEFQVGVGRARGCCTSDNAAAFALINASSKEEALQITIREKIPPAALFOFHNLNSNL	60	
Db	1	tlvgeqfvgdgrsarqscqtssndaaaealnasskeeaqlitirekipekylifqfminsal	60	
QY	61	DRIFDKTPEP	70	
Db	61	aalifdktppep	70	
XX				
XX	RESULT 15			
XX	AAB18682			
XX	AAB18682 standard; peptide: 70 AA.			
XX	AAB18682;			
XX				

Job time: 527 sec

```

DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KM ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key
FH Location/Qualifiers
FT Misc-difference 52 /note= "wild type residue replaced with Ala"
FT Misc-difference 54 /note= "wild type residue replaced with Ala"
FT Misc-difference 55 /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNCL-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI: 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
PS Claim 53: Page 44-45; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA:

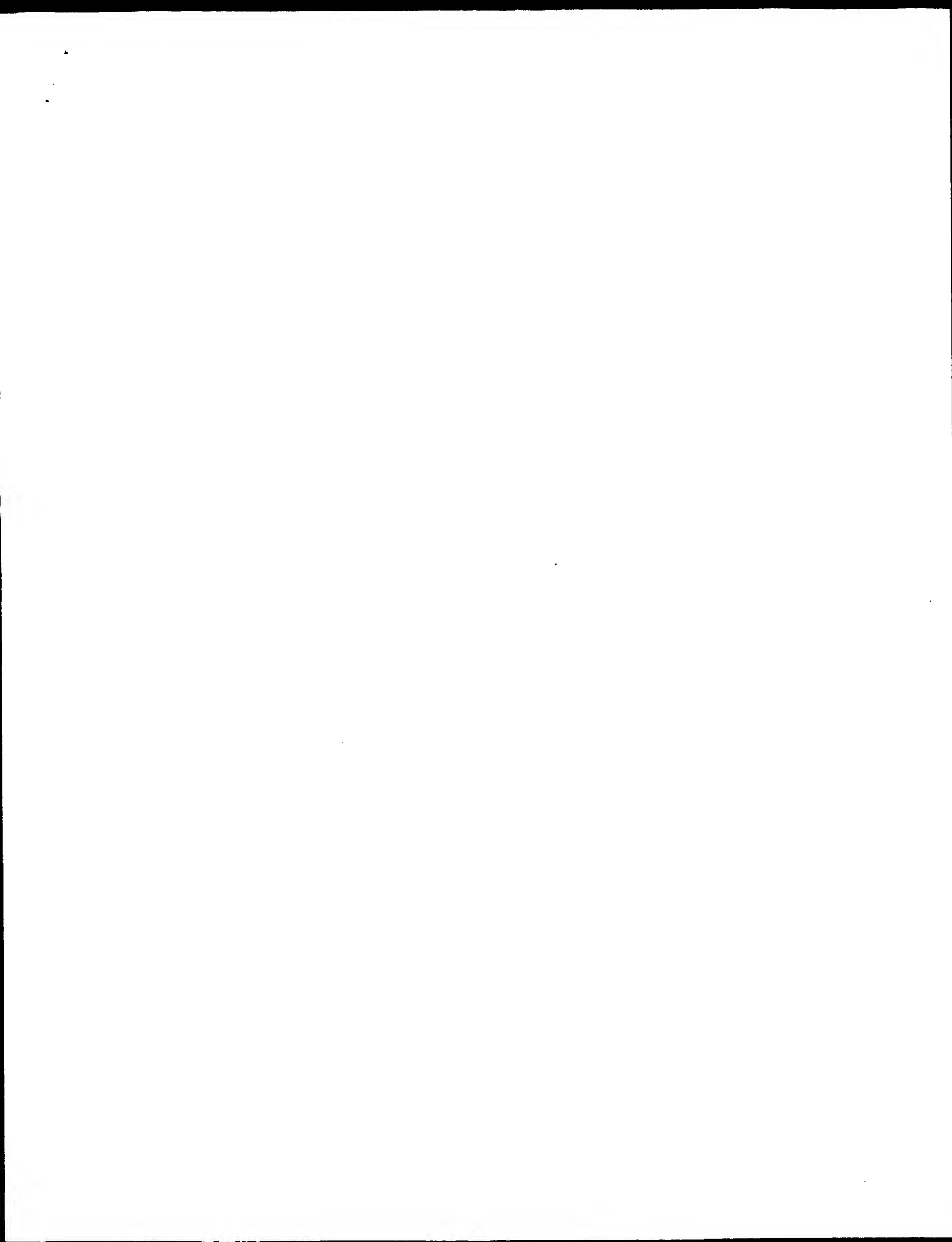
Query Match 88.9%; Score 319; DB 21: Length 70;
Best Local Similarity 91.4%; Pred. No. 5, Re-34:
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLVWGFQVNDGRSAGCGCTSDAAGALNASSKEEALQIRREKIPAAALFQPHNLNSNL 60
   |||||||
Db 1 tlvwgfqvgdgrsargcgctsdnaaalaasskeaalqirrekipekylfafaalnsnl 60
   |||||||

QY 61 DRIFDKTPEP 70
   |||||||
Db 61 drifdktppep 70
   |||||||

```

Search completed: October 10, 2002, 01:47:18



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds  
(Without alignments)  
102.295 Million cell updates/sec

Title: US-09-289-346a-6

Perfect score: 357  
Sequence: 1 TLWGEFQVDRSARGGCGT.....FAFALNSNLDRIKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL.19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	291	81.5	351	12	Q91R10 tomato seve
2	289	81.0	352	12	Q9E000 tomato rugo
3	272	76.2	226	12	009727 leonurus mo
4	272	76.2	226	12	Q9WHF6 tomato mild
5	270	75.6	361	12	067574 bean golden
6	268	75.1	225	12	Q9QDB1 cowpea gold
7	268	75.1	314	12	Q9ELT8 sweet potat
8	268	75.1	364	12	Q9G555 sweet polat
9	265	74.2	185	12	Q98693 sida golden
10	257	72.0	149	12	P88975 macroptilin
11	257	72.0	233	12	Q9YL44 macroptilin
12	251	70.3	234	12	Q91R80 geminivirid
13	248	69.5	190	12	Q9Z089 tobacco lea
14	248	69.5	190	12	Q9Z084 tobacco lea
15	247	69.2	190	12	Q9W827 tobacco lea
16	247	69.2	208	12	Q9Z0C4 tobacco lea

RESULT	ID	Q91R10	PRELIMINARY:	PRT:	351 AA.	ALIGNMENTS
AC	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10	Q91R10
DT	01-DEC-2001	(TREMUREL. 19, Created)				
DT	01-DEC-2001	(TREMUREL. 19, Last sequence update)				
DT	01-DEC-2001	(TREMUREL. 19, Last annotation update)				
DE	REP PROTEIN.					
GN	AC1.					
OS	Tomato severe rugose virus.					
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.					
OX	NCBI_TaxID=158463;					
RN	[1]					
RC	SEQUENCE FROM N.A.					
RA	STRAIN-MINAS GERAIS;					
RT	Reverde W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F., "The full-length DNA-A nucleotide sequence of a novel tomato-infecting RT begomovirus, Tomato severe rugose virus, in Brazil."					
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.					
DR	EMBL: AY029750; AAK50357.1; - FMBL: AY029750; AAK50357.1; -					
SQ	SEQUENCE 351 AA: 40122 MW: 87P937A4F873B6CF CRC64:					

Query Match: 81.5%; Score 291; DB 12; Length 351;  
Best Local Similarity: 78.6%; Pred. No. 2.6e-25;  
Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Query	1	TLWGEFQVDRSARGGCGTSDAAAEALNASSKEEALQITREKIPKYLFAFALNSNL 60
DB	111	TLWGEFQVDRSARGGCGTANDAAAEALNAPSKDALQITREKIPKYLFAFALNSNL 170
QY	61	DRIPKTPPEP 70
DB	171	DRIFARPEP 180
RESULT	2	
Q9E000		PRELIMINARY: PRT: 352 AA.
ID	Q9E000	

AC 09E000;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN.  
 GN ACI.  
 OS Tomato rugose mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=134599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,  
 RA Zambolim E.M., Zerbini E.M.;  
 RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
 RT (TRMV), a Begomovirus isolated from Tomatoes at Triangulo Mineiro,  
 RT Minas Gerais, Brazil.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF291105; AAC15546.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 SQ SEQUENCE 352 AA; 40012 MW; 47CD55838E24D613 CRC64;

Query Match 81.0%; Score 289; DB 12; Length 352;  
 Best Local Similarity 78.6%; Pred. No. 4.3e-25;  
 Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 1 TLVWGEFOVDSARGGCGTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60  
 Db 111 TIEWGEFOVDSARGGCGTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 170  
 OY 61 DRIFDKTPEP 70  
 Db 171 DRIFAKAPEP 180

RESULT 3

ID 009727 PRELIMINARY; PRT; 226 AA.  
 AC 009727;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonurus mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=58177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-LEWY-BRAZIL 1;  
 RA Faria J.C., Maxwell D.P.;  
 RT "Variability in geminivirus associated with Phaseolus vulgaris in  
 RT Brazil.";  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U92572; AA851157.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 FT NON-TER 226  
 SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 76.2%; Score 272; DB 12; Length 226;  
 Best Local Similarity 77.1%; Pred. No. 2.3e-23;  
 Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVWGEFOVDSARGGCGTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60  
 Db 111 TIEWGEFOVDSARGGCGTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 170

OY 61 DRIFDKTPEP 70  
 Db 171 DRIFAKAPEP 180

RESULT 4

ID 09WHF6 PRELIMINARY; PRT; 226 AA.  
 AC 09WHF6;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Tomato mild mottle geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
 OX NCBI\_TaxID=92943;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-HN96-H5KW;  
 RA Nakula M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
 RA Maxwell D.P.;  
 RT "Molecular characterization and DNA-based detection methods for  
 RT vegetable-infecting geminiviruses in Central America.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF131071; AA033471.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 FT NON-TER 226  
 SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 76.2%; Score 272; DB 12; Length 226;  
 Best Local Similarity 72.9%; Pred. No. 2.3e-23;  
 Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

OY 1 TLVWGEFOVDSARGGCGTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60  
 Db 111 TIEWGEFOVDSARGGCGTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 170  
 OY 61 DRIFDKTPEP 70  
 Db 171 DRIFAKAPEP 180

RESULT 5

ID 067574 PRELIMINARY; PRT; 361 AA.  
 AC 067574;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE POTATIVE REPLICATIVE PROTEIN.  
 GN AL1.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlgvist P.G.,  
 RA Maxwell D.P., Russell D.R.;  
 RT "Cloning of the complete DNA genomes of four bean-infecting  
 RT geminiviruses and determining their infectivity by electric discharge  
 RT particle acceleration.";  
 RL Phytopathology 81:980-985(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 RA Morales F.J., Maxwell D.P.;  
 RT "Differentiation of Bean-Infecting geminiviruses by nucleic acid



RT hybridization probes and aspects of bean golden mosaic in Brazil.";  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
 RT "Genetic diversity in geminiviruses causing bean golden mosaic disease: the nucleotide sequence of the infectious cloned DNA components of a Brazilian isolate of bean golden mosaic geminivirus.";  
 RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; M88686; AAA46312.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 SQ SEQUENCE 361 AA: 41041 MW: 0094C7ACAF06B788 CRC64;

Query Match 75.6%; Score 270; DB 12; Length 361;  
 Best Local Similarity 76.1%; Pred. No. 6.8e-23;  
 Matches 51; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 4 WGEFVDSRARGCGCTNSDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNDRT 63  
 DB 113 WGHFVDSRARGCGCTNSDAAAEALNASSKEEALQIIREKIPKYLFOYHNLSSNDRT 172  
 QY 64 FDKTPEP 70  
 DB 173 FTKAPDP 179

RESULT 6  
 Q90DB1 PRELIMINARY; PRT; 225 AA.  
 AC Q90DB1;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=69263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGMV-BR;  
 RA Faria J.C.;  
 RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from Brazil.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF188708; AAF06318.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 FT NON-TER 225  
 SQ SEQUENCE 225 AA: 25766 MW: 1089CB6BD8D15B5D CRC64;

Query Match 75.1%; Score 268; DB 12; Length 225;  
 Best Local Similarity 76.1%; Pred. No. 6.6e-23;  
 Matches 51; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 WGEFVDSRARGCGCTNSDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNDRT 63  
 DB 113 WGHFVDSRARGCGCTNSDAAAEALNASSKEEALQIIREKIPKYLFOYHNLSSNDRT 172  
 QY 64 FDKTPEP 70  
 DB 173 FKKPEP 179

RESULT 7  
 Q9ELT8

ID Q9ELT8 PRELIMINARY; PRT; 314 AA.  
 AC Q9ELT8;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United States.";  
 RT Plant Dis. 82:1253-1257(1998).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A. to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF288227; AAG01006.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 SQ SEQUENCE 314 AA: 35153 MW: 6B6220613046943F CRC64;

Query Match 75.1%; Score 268; DB 12; Length 314;  
 Best Local Similarity 79.4%; Pred. No. 9.8e-23;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLWGEFVDSRARGCGCTNSDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
 DB 110 TITWGEFVDSRARGCGCTNSDAAAEALNASSKEEALQIIREKIPKYLFOYHNLSSNL 169  
 QY 61 DRIETKP 68  
 DB 170 DRIETKP 177

RESULT 8  
 Q9OS55 PRELIMINARY; PRT; 364 AA.  
 AC Q9OS55;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United States.";  
 RT Plant Dis. 82:1253-1257(1998).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl virus: nucleotide sequence and phylogenetic relationships.";  
 RT Online Publication.  
 DR EMBL; AF104036; AAD47173.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 SQ SEQUENCE 364 AA: 40680 MW: 5F79752431A09D6E CRC64;

```

Query Match
Best Local Similarity 75.1%; Score 268; DB 12; Length 364;
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 60
   110 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 169
Db 61 DRIFDKTP 68
   170 DRIFSKPP 177

RESULT 9
Q98693 PRELIMINARY; PRT; 185 AA.
AC Q98693;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN AC1.
OS sida golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51034;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA;
RA Roye M.E.; McLaughlin W.A., Nakha N.K., Maxwell D.P.;
RT "Genetic diversity among geminiviruses associated with the weed
RT species sida spp, Macropitillum lathyroides, and wissadula amplissima
RT from Jamaica."
RL Plant Dis. 81:1251-1258(1997).
DR EMBL; U67926; AAB97865.1; -
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRODOM: PD000736; Gemini_AL1.
FT NON_TER 1
FT NON_TER 185
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match
Best Local Similarity 74.2%; Score 265; DB 12; Length 185;
Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 60
   89 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 148
Db 61 DRIFDKTP 70
   149 DRIFSKPP 158

RESULT 10
P88975 PRELIMINARY; PRT; 149 AA.
AC P88975;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN AC1.
OS Macropitillum golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA;
RA Roye M.E.;
Thesis (1996), Biochemistry, University of the West Indies, Jamaica.

```

```

DR EMBL; U75278; AAB36919.1; -
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRODOM: PD000736; Gemini_AL1.
FT NON_TER 1
FT NON_TER 149
SQ SEQUENCE 149 AA; 16785 MW; EACF5EDAC9CD508 CRC64;

Query Match
Best Local Similarity 72.0%; Score 257; DB 12; Length 149;
Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 60
   52 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 111
Db 61 DRIFDKTP 70
   112 DRIFSKPP 121

RESULT 11
Q9YL44 PRELIMINARY; PRT; 233 AA.
AC Q9YL44;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS Macropitillum golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E.;
RT "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT from Jamaica."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E.; McLaughlin W.A., Maxwell D.P.;
RT "Molecular characterization of two distinct geminiviruses infecting M.
RT lathyroides from Jamaica."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF098940; AAD17850.1; -
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRODOM: PD000736; Gemini_AL1.
FT NON_TER 233
FT NON_TER 233
SQ SEQUENCE 233 AA; 26356 MW; AA490AF4D2166A02 CRC64;

Query Match
Best Local Similarity 72.0%; Score 257; DB 12; Length 233;
Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 60
   110 TLWGEFQVDSRSGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 169
Db 61 DRIFDKTP 70
   170 DRIFSKPP 179

RESULT 12
O39180

```

ID 039180 PRELIMINARY: PRT: 234 AA.  
 AC 039180;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE AC1 (FRAGMENT).  
 OS Geminiiviridae.  
 OC Viruses; ssDNA viruses.  
 OX NCBI\_TaxID=10811;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TOMATO STRAIN;  
 RA Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;  
 RT "Partial Characterization of Two Whitefly-Transmitted Geminiiviruses  
 Infecting Tomatoes in Venezuela.";  
 RL Plant Dis. 81:312-312(1997).  
 DR EMBL: AF026553; AAB82605.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 234  
 SQ SEQUENCE 234 AA: 26486 MW: 9EDBF0697105CD19 CRC64;

Query Match 70.3%; Score 251; DB 12; Length 234;  
 Best Local Similarity 67.1%; Pred. No. 6.2e-21;  
 Matches 47; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDGSRAGCGCOTSDNAAEALNASSKEALQIIREKIPKYLFAFALNSNL 60  
 DB 110 TLEWGLFQIDGRSARSGCGQGTANDAAEALNSGTKEAMKRIKEKLEPEKLFQYHNLSL 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 DRIFPKAPNP 179

RESULT 13  
 ID 092089 PRELIMINARY: PRT: 190 AA.  
 AC 092089;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YOKOHAMA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 190  
 SQ SEQUENCE 190 AA: 21432 MW: AAC093D1D1610FAD CRC64;

Query Match 69.5%; Score 248; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 1.1e-20;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;  
 OY 1 TLVWGEFQVDGSRAGCGCOTSDNAAEALNASSKEALQIIREKIPKYLFAFALNSNL 60  
 DB 110 TLEWGLFQIDGRSARSGCGQGTANDAAEALNSGTKEAMKRIKEKLEPEKLFQYHNLSL 169

DB 85 TLEWGLFQIDGRSARSGCGQGTANDAAEALNASSKEALQIIREKIPKYLFAFALNSNL 144  
 OY 61 DRI-----FDKTPPE 69  
 DB 145 DRIFAPLEFVFCPTASSFDQVPE 169

RESULT 14  
 ID 092084 PRELIMINARY: PRT: 190 AA.  
 AC 092084;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-2, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YOKOHAMA5;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001318; BAA34039.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 190  
 SQ SEQUENCE 190 AA: 21444 MW: AAC1C2943E3F01AD CRC64;

Query Match 69.5%; Score 248; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 1.1e-20;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

OY 1 TLVWGEFQVDGSRAGCGCOTSDNAAEALNASSKEALQIIREKIPKYLFAFALNSNL 60  
 DB 85 TLEWGLFQIDGRSARSGCGQGTANDAAEALNASSKEALQIIREKIPKYLFAFALNSNL 144  
 OY 61 DRI-----FDKTPPE 69  
 DB 145 DRIFAPLEFVFCPTASSFDQVPE 169

RESULT 15  
 ID 09W827 PRELIMINARY: PRT: 190 AA.  
 AC 09W827;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE CORAL-1, PARTIAL AND  
 DE COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CORAL;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001303; BAA34010.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.

FT NON\_TER 1 1  
 FT NON\_TER 190 190  
 SQ SEQUENCE 190 AA: 21444 MW: 93C3742A8EBDB7EB CRC64:

Query Match 69.2%; Score 247; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 1.4e-20;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;  
 QY 1 TLVWGEFQYDGRSARGCQTSNDAAAEALNASSKEEALQITREKIPKYLFAFAALNSNL 60  
 85 TLEMGTFQYDGRSARGCQCNANDACAEALNASSKAALSIIREKLPKDFIFQYHNLNSNL 144  
 QY 61 DRT-----FDKTPF 69  
 145 DRIFADPLEFVCPFSNSSFDQVPE 169  
 Db

Search completed: October 10, 2002, 02:01:59  
 Job time: 1092 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:52 ; Search time 34.62 seconds

(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346a-6

Perfect score: 357  
Sequence: 1 TLWGEFQVDSRGSGCOT.....FAFALNSLDRIDFKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	340	95.2	352	1 VAL1_TGMV	P03567 tomato gold
2	240	67.2	361	1 VAL1_PYMV	P27258 potato yell
3	227	63.6	358	1 VAL1_CLVX	P14982 cassava lat
4	227	63.6	358	1 VAL1_CLVN	P14972 cassava lat
5	220	61.6	362	1 VAL1_TYCA	P36279 tomato yell
6	217	60.8	359	1 VAL1_TYCU	P38609 tomato yell
7	210	58.8	359	1 VAL1_TYLCM	P27260 tomato yell
8	209	58.5	349	1 VAL1_PHOV	Q06923 pepper huas
9	204	57.1	353	1 VAL1_BGMV	P05175 bean golden
10	203	56.9	355	1 VAL1_ABMVW	P21947 abutilon mo
11	200	56.0	358	1 VAL1_BCTV	P14991 beet curly
12	200	56.0	361	1 VAL1_TMOV	Q06567 tomato molt
13	191	53.5	357	1 VAL1_TYCV	P27259 tomato yell
14	194	29.1	347	1 VAL1_SLICV	P29048 squash leaf
15	67	18.8	447	1 TBL2_HUMAN	Q94943 homo sapien
16	64.5	18.1	1713	1 LMA3_HUMAN	Q16787 homo sapien
17	63.5	17.8	1610	1 CCAD_HUMAN	Q09244 mesocricetu
18	63.5	17.8	2161	1 CCAD_HUMAN	Q01668 homo sapien
19	63.5	17.8	2203	1 CCAD_HUMAN	P27732 ratius norv
20	60	16.8	269	1 ORN_YEAST	P4964 saccharomyc
21	60	16.8	295	1 VAL1_TYDVA	P16167 tobacco yell
22	60	16.8	355	1 CRTB_RHOSH	P54905 rhodobacter
23	59.5	16.7	299	1 Y175_HELPJ	Q92567 helicobacte
24	59.5	16.7	470	1 RHB4_RHIME	Q92382 rhizobium m
25	59.5	16.7	4385	1 YP73_CABEL	Q09222 caenorhabdi
26	58.5	16.4	706	1 Y006_RICPR	Q9264 caenorhabdi
27	58.5	16.4	1232	1 Y005_CABEL	P34643 caenorhabdi
28	58.5	16.2	207	1 IL6_MARMO	O35736 marioneta mon
29	58	16.2	319	1 IDH2_STAAM	O99135 staphylococ
30	58	16.2	447	1 CDS4_DROME	P56079 d phosphati
31	57.5	16.1	367	1 LHX4_MOUSE	P53776 mus musculi
32	57.5	16.1	844	1 SECA_STACA	P47994 staphylococ
33	57.5	16.1	1852	1 CCAS_CYPCA	P22316 cyprius ca

34	57	16.0	1483	1 CYP1_YEAST	P12351 saccharomyc
35	57	16.0	1608	1 HLYA_SERMA	P15320 serralia ma
36	56.5	15.8	511	1 HUTH_VIRCH	O9K94 vibrio chol
37	56.5	15.8	2190	1 CCAD_CHICK	O73700 gallus gall
38	56	15.7	258	1 YL48_MYCTO	O06228 mycobacteri
39	56	15.7	408	1 CINA_THEMA	Q955x1 thermotoga
40	56	15.7	492	1 MOT3_MOUSE	O35308 mus musculi
41	56	15.7	492	1 MOT3_MOUSE	O70461 ratius norv
42	56	15.7	703	1 GYS2_RAT	P17625 ratius norv
43	56	15.7	767	1 TOP1_MOUSE	O04750 mus musculi
44	56	15.7	1033	1 YDK9_SCHPO	P87115 schizosacch
45	56	15.7	1051	1 CARR_SULSO	O59969 sulfobus

## ALIGNMENTS

RESULT	ID	VAL1_TGMV	STANDARD	PRT	352 AA
AC	P03567				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	01-JUN-1994	(Rel. 29, Last annotation update)			
DE	AL1 protein.				
GN	AC1.				
OS	Tomato golden mosaic virus (TGMV).				
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.				
OX	NCBI_Taxid=10831;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck R.W.;				
RT	"Complete nucleotide sequence of the infectious cloned DNA components				
RT	of tomato golden mosaic virus: potential coding regions and regulatory				
RT	sequences."				
RL	EMBO J. 3:2197-2205(1984).				
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL: K02029; ; NOT_ANNOTATED_CDS.				
DR	PIR: A04170; OQCVL1.				
DR	InterPro: IPR001191; Gemin1_AL1.				
DR	Pfam: PF00799; Gemin1_AL1.1.				
DR	PRINTS: PR00227; GEMCOATL1.				
DR	ProDom: PD000736; Gemin1_AL1.1.				
KW	ATP-binding; 223				
FT	NP-BIND				
FT	SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;				
QY	1 TLWGEFQVDSRGSGCOTSDAAAEALNASSKEEALQIRREIKYLFAPALNSNL 60				
QY					
Db	111 TLWGEFQVDSRGSGCOTSDAAAEALNASSKEEALQIRREIKYLFAPALNSNL 170				
QY	61 DRIFDKTPPEP 70				
QY					
Db	171 DRIFDKTPPEP 180				
RESULT	2				
VAL1_PYMV					
ID	VAL1_PYMV	STANDARD;	PRT;	361 AA.	

Query Match 95.2%; Score 340; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 1.2e-30;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 protein.
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311403; PubMed=1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RT potato yellow mosaic virus."
RL J. Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00940; BAA00782.1; -
DR PIR: J00364; Q0CVPF.
DR InterPro: IPR001191; Geminl_AL1.
DR Pfam: PF00799; Geminl_AL1; 1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Geminl_AL1; 1.
KW ATP-binding.
FT NP_BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match 67.2%; Score 240; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 1.7e-19;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNMSKEEALQIREKPEKYLFAFAALNSNL 60
DB 110 TLVWGEFQIDGRSARGGQGVNDAAAEALNMSGTKEAAKTIKTKLEKFLFYHNLSNL 169
QY 61 DRIFDKTPEP 69
DB 170 DRIFDKAPE 178

RESULT 3
VAL1_CLVK STANDARD; PRT; 358 AA.
ID VAL1_CLVK
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RX Stanley J., Gay W.R.;
RT "Nucleotide sequence of cassava latent virus DNA."
RT Nature 301:260-262(1983).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02057; -; NOT_ANNOTATED_CDS.
DR InterPro: IPR001191; Geminl_AL1.
DR Pfam: PF00799; Geminl_AL1; 1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Geminl_AL1; 1.
KW ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 63.6%; Score 227; DB 1; Length 358;
Best Local Similarity 60.0%; Pred. No. 4.8e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNMSKEEALQIREKPEKYLFAFAALNSNL 60
DB 109 TLVWGEFQIDGRSARGGQGSANDAYAKALNMSGSEALNVIKDFVLQFHNLSNL 168
QY 61 DRIFDKTPEP 70
DB 169 DRIFDEPPAP 178

RESULT 4
VAL1_CLVK STANDARD; PRT; 358 AA.
ID VAL1_CLVK
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain)".
RT Nucleic Acids Res. 18:197-198(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17095; CAA34953.1; -
DR PIR: S07594;
DR InterPro: IPR001191; Geminl_AL1.
DR Pfam: PF00799; Geminl_AL1; 1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Geminl_AL1; 1.
KW ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40435 MW; 1DB16BB0CB2D5E2C CRC64;

Query Match 63.6%; Score 227; DB 1; Length 358;
Best Local Similarity 60.0%; Pred. No. 4.8e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNMSKEEALQIREKPEKYLFAFAALNSNL 60
DB 109 TLVWGEFQIDGRSARGGQGSANDAYAKALNMSGSEALNVIKDFVLQFHNLSNL 168

```

OY 61 DRFDKTEP 70  
 |||| : ||  
 DB 169 DRFDKTEP 178

## RESULT 5

VAL1\_TYLCU STANDARD: PRT: 362 AA.  
 AC P36279;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE A11 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36447;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93139778; PubMed=8423446;  
 RA Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;  
 RT "Nucleotide sequence and genome organization of tomato leaf curl  
 geminivirus.";  
 RT J. Gen. Virol. 74:147-151(1993).  
 RL -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.  
 DR InterPro: IPR001191; Gemin1\_ALL.  
 DR Pfam: PF00799; Gemin1\_ALL.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemin1\_ALL.1.  
 KW ATP-binding.  
 FT NP\_BIND 221 228 ATP (POTENTIAL).  
 SO SEQUENCE 362 AA; 41197 MW; 343E7184B4704098 CRC64;

Query Match 61.6%; Score 220; DB 1; Length 362;  
 Best Local Similarity 52.9%; Pred. No. 3e-17;  
 Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

OY 1 TLWGEFQVGRSARGCGQTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSND 60  
 || |||||:|||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 DB 110 TLEWGEFOIDGRSARGGQOANDAYAKALNAGSKSEALDVYKELAPDYTLHFHNSND 169  
 || |||||:|||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 OY 61 DRFDKTEP 69  
 |||| : ||  
 DB 170 DRFDKTEP 194

## RESULT 6

VAL1\_TYLCU STANDARD: PRT: 359 AA.  
 AC P38609;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE A11 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=37139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94256836; PubMed=8198442;  
 RA Norris E., Hidalgo E., Accotto G., Moriones E.;  
 RT "High similarity among the tomato yellow leaf curl virus isolates  
 from the west Mediterranean basin: the nucleotide sequence of an  
 infectious clone from Spain.";  
 RT Arch. Virol. 135:165-170(1994).  
 RL Arch. Virol. 135:165-170(1994).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z25751; CAA81026.1; -;  
 DR PIR: S39211; S39211.  
 DR InterPro: IPR001191; Gemin1\_ALL.  
 DR Pfam: PF00799; Gemin1\_ALL.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemin1\_ALL.1.  
 KW ATP-binding.  
 FT NP\_BIND 221 228 ATP (POTENTIAL).  
 SO SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 60.8%; Score 217; DB 1; Length 359;  
 Best Local Similarity 59.4%; Pred. No. 6.3e-17;  
 Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

OY 2 TLWGEFQVGRSARGCGQTNDAAAEALNASSKEEALQIIRKIPKYLFAFALNSND 61  
 || |||||:|||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 DB 111 TLEWGEFOIDGRSARGGQOANDAYAKALNAGSKSEALDVYKELAPDYTLHFHNSND 170  
 || |||||:|||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 OY 62 DRFDKTEP 70  
 |||| : ||  
 DB 171 DRFDKTEP 179

## RESULT 7

VAL1\_TYLCU STANDARD: PRT: 359 AA.  
 AC P27260;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A11 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92107660; PubMed=1840676;  
 RA Rheyf-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,  
 RA Gronenborn B.;  
 RT "Tomato yellow leaf curl virus from Sardinia is a  
 RT whitefly-transmitted monopartite geminivirus.";  
 RL Nucleic Acids Res. 19:6763-6769(1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X61153; CAA3466.1; -;  
 DR PIR: S22593; S22593.  
 DR InterPro: IPR001191; Gemin1\_ALL.  
 DR Pfam: PF00799; Gemin1\_ALL.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemin1\_ALL.1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SO SEQUENCE 359 AA; 40733 MW; 9717B4A07C93FEA7 CRC64;

Query Match 58.8%; Score 210; DB 1; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 3.8e-16;

Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDGRSARGCQTSNDAAAEALNASSKEEALQIREKIPKYLFAFALNSNL 61  
 Db 111 LEMGTFOIDGRSARGGQGTANDAYAKAINAGSKSQALDIKELAPROYLVLFPHNINSNL 170

QY 62 RIFDKTPPEP 70  
 Db 171 KVFQVPPAP 179

RESULT 8  
 VALI\_PHV STANDARD; PRT; 349 AA.

AC 006923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN AL1.  
 OS Peppier huasteco virus (PHV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=28349;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94015007; PubMed=8409944;  
 RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,  
 RA Rivera-Bustamante R.F.;  
 RT "Complete nucleotide sequence of pepper huasteco virus: analysis and  
 comparison with bipartite geminiviruses.";  
 RL J. Gen. Virol. 74:2225-2231(1993).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X70418; CAA49856.1; -  
 DR PIR: S31875; S31875.  
 DR PIR: JQ2300; JQ2300.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND 221 228 ATP (BY SIMILARITY).  
 SQ SEQUENCE 349 AA; 39722 MW; D5F4E76CD56370FA CRC64;

Query Match 58.5%; Score 209; DB 1; Length 349;  
 Best Local Similarity 57.1%; Pred No. 4.8e-16;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGCQTSNDAAAEALNASSKEEALQIREKIPKYLFAFALNSNL 60  
 Db 110 TLVWGEFQVDGRSARGGQGTANDAYAKALNSASAEALQIREKIPKYLFAFALNSNA 169

QY 61 DRIFDKTPPEP 70  
 Db 170 NRIFQTPPEP 179

RESULT 9  
 VALI\_BGMV STANDARD; PRT; 353 AA.

AC P05175;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE AL1 protein (40.2 kDa protein).  
 GN AC1.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;  
 RT "Nucleotide sequence of bean golden mosaic virus and a model for gene  
 regulation in geminiviruses.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M10070; AAA46318.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND 222 229 ATP (POTENTIAL).  
 SQ SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match 57.1%; Score 204; DB 1; Length 353;  
 Best Local Similarity 57.1%; Pred No. 1.8e-15;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGCQTSNDAAAEALNASSKEEALQIREKIPKYLFAFALNSNL 60  
 Db 110 TLVWGEFQVDGRSARGGQGTANDAYAKALNSASAEALQIREKIPKYLFAFALNSNA 169

QY 61 DRIFDKTPPEP 70  
 Db 170 NRIFQTPPEP 179

RESULT 10  
 VALI\_ABMV STANDARD; PRT; 355 AA.

AC P21947;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN AC1.  
 OS Abutilon mosaic virus (isolate West India).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91020984; PubMed=2219703;  
 RA Frischmuth T., Zimat G., Jeske H.;  
 RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
 as well as eukaryotic features.";  
 RL Virology 178:461-468(1990).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



DR EMBL: X15983; NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; OOCVW1.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 DR ATP-binding. 221 228 ATP (POTENTIAL).  
 FT NP\_BIND  
 SQ SEQUENCE 355 AA: 40257 MW: 16A2CA8A63251E95 CRC64;

Query Match 56.9%; Score 203; DB 1; Length 355;  
 Best Local Similarity 55.7%; Pred. No. 2.3e-15;  
 Matches 39; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 1 TLWGGEFQVGRSARCGCCTSDNAAEALNASSKEEALQIREKPEKYLFAFAALNSNL 60  
 DB 110 TLWGGEFQIDGRSARCGQGTANDSYAKALNAGDVQSLNLIKEDPKDYLVONHNIRSNL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 11  
 VAL1\_BCTV STANDARD; PRT: 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE AL1 protein (40.8 kDa protein).  
 OS Beet curly top virus (BCTV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
 OX NCBI\_TaxID=10840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Plummer M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 RT beet curly top virus."  
 RL EMBO J. 5:1761-1767(1986).  
 CC - SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: X04144; NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 KW ATP-binding. 222 229 ATP (POTENTIAL).  
 FT NP\_BIND  
 SQ SEQUENCE 358 AA: 40889 MW: 39A45F3C0B9C333 CRC64;

Query Match 56.0%; Score 200; DB 1; Length 358;  
 Best Local Similarity 52.9%; Pred. No. 5e-15;  
 Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLWGGEFQVGRSARCGCCTSDNAAEALNASSKEEALQIREKPEKYLFAFAALNSNL 60  
 DB 110 TLWGGEFQIDGRSARCGQGTANDSYAKALNAGDVQSLNLIKEDPKDYLVONHNIRSNL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 QKIFORPPDP 179

RESULT 12  
 VAL1\_TMOV STANDARD; PRT: 361 AA.  
 AC Q06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN AL1.  
 OS Tomato mottle virus (isolate Florida) (TMOV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93107858; PubMed=1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 RT isolated from tomatoes in Florida."  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC - SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: L14460; AAC32414.1;  
 DR PIR: J01870; J01870.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.  
 KW ATP-binding. 222 229 ATP (BY SIMILARITY).  
 FT NP\_BIND  
 SQ SEQUENCE 361 AA: 40516 MW: 8138B65CEAC6950 CRC64;

Query Match 56.0%; Score 200; DB 1; Length 361;  
 Best Local Similarity 52.9%; Pred. No. 5e-15;  
 Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLWGGEFQVGRSARCGCCTSDNAAEALNASSKEEALQIREKPEKYLFAFAALNSNL 60  
 DB 110 TLWGGEFQIDGRSARCGQGTANDSYAKALNAGDVQSLNLIKEDPKDYLVONHNIRSNL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT: 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RT with a single genomic component."  
 RL Virology 185:151-161(1991).

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: X15656; CAA33688.1; -  
 DR PIR: D40779; QOCVCL.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP-BIND. 219 226 ATP (POTENTIAL).  
 SQ SEQUENCE 357 AA; 40678 MW; 939AB6E1AB3E2A7 CRC64;  
 Query Match 53.5%; Score 191; DB 1; Length 357;  
 Best Local Similarity 62.3%; Pred. No. 5e-14;  
 Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;  
 QY 4 GEFQVDGRSARCGCQTSNDAAEALNASSKEBALQIREKIPEKYLFAFALNSLDRIF 63  
 DB 111 FGVSQIDGRSARCGQGSANDAYAEALNSGSKSEALNITLKEKAPKDYLLQHNLSNDR1 170  
 QY 64 F 64  
 DB 171 F 171  
 RESULT 14  
 VAL1\_SLCV STANDARD; PRT; 347 AA.  
 ID VAL1\_SLCV  
 AC P29048;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE AL1 protein.  
 OS Squash leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 ON NCBI\_TaxID=10829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91082449; PubMed=1984668;  
 RA Lazarowitz S.G., Lazdins I.B.;  
 RT "Infectivity and complete nucleotide sequence of the cloned genomic  
 RT components of a bipartite squash leaf curl geminivirus with a broad  
 RT host range phenotype.";  
 RL Virology 180:58-69(1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: M38183; AAC32410.1; ALT\_INIT.  
 DR PIR: C36785; QOCVSL.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP-BIND. 218 225 ATP (POTENTIAL).  
 SQ SEQUENCE 347 AA; 39110 MW; AFDABDEDE122110E CRC64;

Query Match 29.1%; Score 104; DB 1; Length 347;  
 Best Local Similarity 34.8%; Pred. No. 0.00024;  
 Matches 23; Conservative 14; Mismatches 25; Indels 4; Gaps 1;  
 QY 5 GEFQVDGRSARCGCQTSNDAAEALNASSKEBALQIREKIPEKYLFAFALNSLDRIF 64  
 DB 116 GQYKVG---GSKSNKDDVYHNNAVAGSAGCALDIIRKAGDKPTIVYHNLAVERTLF 171  
 QY 65 DKTPRP 70  
 DB 172 QKPEP 177  
 RESULT 15  
 TBL2\_HUMAN STANDARD; PRT; 447 AA.  
 ID TBL2\_HUMAN  
 AC O914E3; O910E2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transducin beta-like 2 protein (WS beta-transducin repeats protein)  
 DE (WS-betaTRP).  
 GN TBL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20044626; PubMed=10575226;  
 RA Perez Jurado L.A., Wang Y.-K., Francke U., Cruces J.;  
 RT "TBL2, a novel transducin family member in the WS deletion:  
 RT characterization of the complete sequence, genomic structure,  
 RT transcripional variants and the mouse ortholog.";  
 RL Cytogenet. Cell Genet. 86:277-284(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glass S.,  
 RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
 RA Mewes H.-W., Oltewald B., Obermaier B., Tampe J., Heubner D.,  
 RA Wambutt R., Korn B., Klein M., Poustka A.;  
 RT "Towards a catalog of human genes and proteins: sequencing and  
 RT analysis of 500 novel complete protein coding human cDNAs.";  
 RL Genome Res. 11:422-435(2001).  
 RN [3]  
 RP SEQUENCE OF 14-447 FROM N.A.  
 RX MEDLINE=99075645; PubMed=9860302;  
 RA Meng X., Lu X., Li Z., Green E.D., Massa H., Trask B.J., Morris C.A.,  
 RA Keating M.T.;  
 RT "Complete physical map of the common deletion region in Williams  
 RT syndrome and identification and characterization of three novel  
 RT genes.";  
 RL Hum. Genet. 103:590-599(1998).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AF097484; AAF06823.1; -  
 DR PIR: AF097485; AAF06824.1; -  
 DR EMBL: AL080162; CAB45751.1; -  
 DR EMBL: AF056183; AAD28083.1; ALT\_INIT.  
 DR MIM: 605842; -  
 DR InterPro: IPR001680; WD40.

DR Pfam: PF00400; WD40; 5.  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR SMART: SM00320; WD40; 5.  
DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
DR PROSITE: PS00682; WD\_REPEATS\_2; 3.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
FT REPEAT 88 127 WD 1.  
FT REPEAT 134 174 WD 2.  
FT REPEAT 186 226 WD 3.  
FT REPEAT 228 267 WD 4.  
FT REPEAT 277 316 WD 5.  
FT REPEAT 329 367 WD 6.  
FT REPEAT 371 409 WD 7.  
SQ SEQUENCE 447 AA: 49797 MW: B26087E1A71D3F9 CRC64;

Query Match 18.8%; Score 67; DB 1; Length 447;  
Best Local Similarity 32.1%; Pred. No. 4.3;  
Matches 18; Conservative 11; Mismatches 23; Indels 4; Gaps 1;  
OY 12 RSARGGQGTNDAAAEALNASSKEEALQITREKIPKYLFAF----AALNSNLDRI 63  
DB 38 RSGRPACQKANGPPDPKSSGSKKQYQRIKKEKPOQHNFTHRLAALAKSHSGNI 93

Search completed: October 10, 2002, 02:09:52  
Job time: 1350 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:44 ; Search time 69.63 Seconds  
(without alignments)  
96,600 Million cell updates/sec

Title: US-09-289-346A-6  
Perfect score: 357  
Sequence: 1 TLWGEFQVYDGRSARGCGQT.....FAFAALNSNLDRIPTKPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 71: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	340	95.2	352	1	QOCVLI
2	240	67.2	361	1	QOCVPT
3	227	63.6	358	2	S07594
4	220	61.6	362	1	J01887
5	217	60.8	359	2	S39211
6	212	59.4	351	2	J02327
7	210	58.8	359	2	S22593
8	209	58.5	349	2	J02300
9	209	58.5	349	2	S31875
10	204	57.1	359	2	S39235
11	203	56.9	355	1	QOCVW1
12	200	56.0	358	1	J01870
13	200	56.0	385	2	S28360
14	193	54.1	360	2	S39885
15	193	53.5	357	1	QOCVCI
16	118	33.1	131	2	S45059
17	104	29.1	347	1	QOCVSI
18	104	28.8	447	2	T12544
19	67	18.1	587	2	J01419
20	64.5	18.1	1713	2	A55347
21	63.5	17.8	180	2	D84082
22	63.5	17.8	1610	2	A46227
23	63.5	17.8	1646	2	JH0422
24	63.5	17.8	2161	2	JH0564
25	63.5	17.8	2181	2	A38198
26	63.5	17.8	2203	2	T43742
27	63	17.6	316	2	C82085
28	62.5	17.5	1033	2	E97700
29	62	17.4	338	2	AG3582

30	62	17.4	340	2	B97011
31	61.5	17.2	392	2	T45290
32	60.5	16.9	840	2	T36175
33	60	16.8	154	2	AD5475
34	60	16.8	269	2	S61633
35	60	16.8	295	2	D42452
36	60	16.8	355	2	S49621
37	60	16.8	355	2	T50746
38	60	16.8	714	2	C95382
39	60	16.8	827	2	D96907
40	60	16.8	1888	2	T14273
41	59.5	16.7	299	2	B71967
42	59.5	16.7	470	2	T46814
43	59.5	16.7	470	2	B95419
44	59.5	16.7	2137	2	T05244
45	59.5	16.7	4385	2	T29042

## ALIGNMENTS

## RESULT 1

QOCVLI  
AL1 protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C>Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
C:Accession: A04170  
R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBO J. 3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <RAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.2% Score 340, DB 1; Length 352;  
Best Local Similarity 95.7% Pred. No. 4.6e-31;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Oy 1 TLWGEFQVYDGRSARGCGQTSNDAAAEALNSKEEALQIRKIPKYLFAFAALNSNL 60  
Db 111 TLWGEFQVYDGRSARGCGQTSNDAAAEALNSKEEALQIRKIPKYLFAFAALNSNL 170  
Oy 61 DRIFDKTPEP 70  
Db 171 DRIFDKTPEP 180  
RESULT 2  
QOCVPT  
AL1 protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUID:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-361 <COO>  
A:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

RESULT 7  
S22593  
hypothetical protein C4 - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
 C:Accession: S22593  
 R:Khey-Pour, A.; Bendahmane, M.; Matzelt, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.  
 Nucleic Acids Res. 19, 6763-6769, 1991  
 A>Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite  
 A:Reference number: S22588; MID:92107660  
 A:Accession: S22593  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-359 <KHE>  
 A:Cross-references: EMBL:X61153; NID:962211; PIDN:CAA43466.1; PID:962217  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 58.8%; Score 210; DB 2; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 3.2e-16;  
 Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVMGEFQVDSARGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 61  
 DB 111 LEMGTFOIDGRSARGCGQTANDAYAKAINAGSKSQALDVKELAPRDYVLFHFNINSND 170  
 OY 62 RIFDKTPPEP 70  
 DB 171 KVFQVPPAP 179

RESULT 8  
 JO2300  
 replicase - pepper huasteco virus (component A)  
 N:Alternate names: ORF AL1 protein  
 C:Species: pepper huasteco virus  
 C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999  
 C:Accession: JO2300  
 R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante, R.F.  
 J. Gen. Virol. 74, 2225-2231, 1993  
 A>Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with bipa  
 A:Reference number: JO2299; MID:94015007  
 A:Accession: JO2300  
 A:Molecule type: DNA  
 A:Residues: 1-349 <TOR>  
 A:Cross-references: GB:X70418; NID:961023; PIDN:CAA49856.1; PID:961025  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 58.5%; Score 209; DB 2; Length 349;  
 Best Local Similarity 57.1%; Pred. No. 4.1e-16;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLWGEFQVDSARGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 60  
 DB 110 TWMGEFQIDGRSARGCGQSANDTYAKALNSASAEALQIIRKPOPHFFLOPHNIVSNA 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 NRIFQTPPEP 179

RESULT 9  
 S31875  
 AL1 protein - pepper rizado amarillo virus  
 C:Species: pepper rizado amarillo virus  
 C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999  
 C:Accession: S31875  
 R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu  
 submitted to the EMBL Data Library, February 1993  
 A>Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa  
 A:Reference number: S31875  
 A:Accession: S31875  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-349 <TOR>

A:Cross-references: EMBL:X70418; NID:961023; PIDN:CAA49856.1; PID:961025  
 A>Note: the source is designated as pepper huasteco virus  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 58.5%; Score 209; DB 2; Length 349;  
 Best Local Similarity 57.1%; Pred. No. 4.1e-16;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLWGEFQVDSARGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 60  
 DB 110 TWMGEFQIDGRSARGCGQSANDTYAKALNSASAEALQIIRKPOPHFFLOPHNIVSNA 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 NRIFQTPPEP 179

RESULT 10  
 S39235  
 gene C1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
 C:Accession: S39235  
 R:Crespi, S.; Norris, B.; Valira, A.; Bosco, D.; Accotto, G.  
 submitted to the EMBL Data Library, December 1993  
 A>Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.  
 A:Reference number: S39233  
 A:Accession: S39235  
 A>Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <CRE>  
 A:Cross-references: EMBL:Z28390; NID:91041671; PID:91334964  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.1%; Score 204; DB 2; Length 359;  
 Best Local Similarity 55.1%; Pred. No. 1.6e-15;  
 Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVMGEFQVDSARGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 61  
 DB 111 LEMGTFOIDGRSARGCGQTANDAYAKAINAGSKSQALDVKELAPRDYVLFHFNINSND 170  
 OY 62 RIFDKTPPEP 70  
 DB 171 KVFQVPPAP 179

RESULT 11  
 OOCVMI  
 AL1 protein - abutilon mosaic virus (isolate West India)  
 C:Species: abutilon mosaic virus  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994  
 C:Accession: A36214  
 R:Frishmuth, T.; Zimmat, G.; Jeske, H.  
 virology 178, 461-468, 1990  
 A>Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as  
 A:Reference number: A36214; MID:91020984  
 A:Accession: A36214  
 A:Molecule type: DNA  
 A:Residues: 1-355 <FRI>  
 A:Cross-references: EMBL:X15983  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 56.9%; Score 203; DB 1; Length 355;  
 Best Local Similarity 55.7%; Pred. No. 2e-15;  
 Matches 39; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

OY 1 TLWGEFQVDSARGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLFAFAALNSNL 60

Db 110 TAEWGEFQIDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 ERIFAKAP 179

## RESULT 12

QJ01870  
 AL1 protein - tomato mottle virus (isolate Florida)  
 C:Species: tomato mottle virus  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01870  
 R:Abouzaid, A.M.; Polston, J.E.; Hiebert, E.  
 J. Gen. Virol. 73, 3225-3229, 1992  
 A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from  
 A:Reference number: J01869; MUID:93107858  
 A:Accession: J01870  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <ABO>  
 A:Cross-references: GB:114460  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 56.0%; Score 200; DB 1; Length 358;  
 Best Local Similarity 52.9%; Pred. No. 4.5e-15;  
 Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
 Db 107 TIEWGEFQIDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 166  
 QY 61 DRIFDKTPEP 70  
 Db 167 ERIFAKAP 176

## RESULT 13

S28360  
 AL1 protein - beet curly top virus  
 C:Species: beet curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 EMBO J. 5, 1761-1767, 1986  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to  
 A:Reference number: S28360  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 56.0%; Score 200; DB 2; Length 385;  
 Best Local Similarity 52.9%; Pred. No. 4.8e-15;  
 Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
 Db 137 TIEWGEFQIDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 196  
 QY 61 DRIFDKTPEP 70  
 Db 197 QKIFORPPDP 206

## RESULT 14

S59885  
 replication-associated protein C1 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 submitted to the EMBL Data Library, February 1995  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
 d geminiviruses.  
 A:Reference number: S58346  
 A:Accession: S59885  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-360 <HON>  
 A:Cross-references: EMBL:248182; NID:944838; PIDN:CAA88229.1; PID:9974211  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 54.1%; Score 193; DB 2; Length 360;  
 Best Local Similarity 57.6%; Pred. No. 2.8e-14;  
 Matches 38; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 4 WGEFQVDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNLDRI 63  
 Db 113 FGVFQIDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNLDRI 172  
 QY 64 FDKTPE 69  
 Db 173 FTFPSAE 178

## RESULT 15

QOCVC1  
 AL1 protein - tomato yellow leaf curl virus  
 N:Alternate names: C1 protein  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 A:Reference number: A40779; MUID:92024070  
 A:Accession: D40779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 A:Cross-references: GB:X15656; NID:962204; PIDN:CAA33688.1; PID:962207  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 53.5%; Score 191; DB 1; Length 357;  
 Best Local Similarity 62.3%; Pred. No. 4.7e-14;  
 Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGEFQVDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNLDRI 63  
 Db 111 FGVFQIDGRSARGGQOTSDNDAAEALNASSKEEALQIIREKIPKYLFAFAALNSNLDRI 170  
 QY 64 F 64  
 Db 171 F 171

Search completed: October 10, 2002, 01:48:44  
 Job time: 562 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds

(without alignments)  
58.394 Million cell updates/sec

Title: US-09-289-346a-6

Perfect score: 357

Sequence: 1 TLVWGEFQVDSRANGSGCQT.....FAFALNSNLDRIFDKTPPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357	100.0	70	21	AA18682
2	340	95.2	70	21	AA18677
3	340	95.2	70	21	AA18687
4	332	93.0	70	21	AA18685
5	329	92.2	70	21	AA18688
6	328	91.9	70	21	AA18692
7	326	91.3	70	21	AA18684
8	326	91.3	70	21	AA18690
9	325	91.0	70	21	AA18678
10	325	91.0	70	21	AA18686
11	324	90.8	70	21	AA18689

12	322	90.2	70	21	AA18680	Mutant peptide der
13	322	90.2	70	21	AA18691	Mutant peptide der
14	319	89.4	70	21	AA18681	Mutant peptide der
15	318	89.1	70	21	AA18683	Mutant peptide der
16	308	86.3	70	21	AA18679	Mutant peptide der
17	210	58.8	359	17	AA18687	Sardinian tomato y
18	210	58.8	359	17	AA18687	Sardinian tomato y
19	210	58.8	359	17	AA18687	Sardinian tomato y
20	206	57.7	353	18	AA18682	Bean golden mosaic
21	206	57.7	353	18	AA18682	Bean golden mosaic
22	206	57.7	353	18	AA18682	Bean golden mosaic
23	206	57.7	353	18	AA18682	Bean golden mosaic
24	206	57.7	353	18	AA18682	Bean golden mosaic
25	204	57.1	353	8	AA18682	ORF 4 gene product
26	204	57.1	361	18	AA18682	Tomato mottle viru
27	204	57.1	361	18	AA18682	Tomato mottle viru
28	204	57.1	361	18	AA18682	Tomato mottle viru
29	204	57.1	361	18	AA18682	Tomato mottle viru
30	197.5	55.3	361	18	AA18682	Product of ORF 4 f
31	192	53.8	362	19	AA18682	Tobacco leaf curl
32	191	53.5	357	18	AA18682	Tomato yellow leaf
33	191	53.5	357	18	AA18682	Tomato yellow leaf
34	191	53.5	357	18	AA18682	Tomato yellow leaf
35	183	51.3	357	18	AA18682	Tomato yellow leaf
36	76.5	21.4	665	22	AA18682	Drosophila melanog
37	67	18.8	447	21	AA18682	Human CRPase assoc
38	67	18.8	447	21	AA18682	Human secreted pro
39	67	18.8	447	21	AA18682	Membrane-bound pro
40	67	18.8	447	21	AA18682	Human adult aorta
41	67	18.8	447	22	AA18682	Human PRO polypept
42	67	18.8	447	22	AA18682	Human polypeptide
43	67	18.8	447	22	AA18682	Human gene 26 enco
44	67	18.8	447	22	AA18682	Human PRO125 (UNQ
45	67	18.8	456	21	AA18682	Human secreted pro

## ALIGNMENTS

RESULT 1	AA18682	standard; peptide; 70 AA.
ID	AA18682	
XX	AA18682	
AC	AA18682	
XX	AA18682	
DT	22-JAN-2001	(first entry)
XX		
DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.	
XX		
KW	Geminivirus: replication protein; Rep protein; AL1: transgenic plant;	
KW	ribosome binding region; resistance; geminivirus infection.	
XX		
OS	Synthetic.	
OS	Tomato golden mosaic virus.	
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 52	/note= "wild type residue replaced with Ala"
FT	Misc-difference 54	/note= "wild type residue replaced with Ala"
FT	Misc-difference 55	/note= "wild type residue replaced with Ala"
FT		
PN	WO200054573-A1.	
XX		
PD	21-SEP-2000.	
XX		
PF	15-MAR-2000; 2000MO-US06759.	
XX		
PR	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX		
PA	(UUNC-) UNIV NORTH CAROLINA STATE.	

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 XX  
 PS Claim 53; Page 44-45; 73pp; English.  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX Sequence 70 AA;

Query Match 100.0%; Score 357; DB 21; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-38;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRSGRGCGTSDNDAAEALNASSKEEAQITREKIPERYLFAFALNSNL 60  
 Db 1 TLVWGEFQVDSRSGRGCGTSDNDAAEALNASSKEEAQITREKIPERYLFAFALNSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 2  
 AAB18677  
 ID AAB18677 standard; peptide: 70 AA.  
 AC AAB18677;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 PN WO200054573-A1.  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI: 2000-618851/59.  
 XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 XX  
 PS Disclosure; Page 18; 73pp; English.  
 CC The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 CC  
 XX Sequence 70 AA;

Query Match 95.2%; Score 340; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 1.3e-35;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRSGRGCGTSDNDAAEALNASSKEEAQITREKIPERYLFAFALNSNL 60  
 Db 1 TLVWGEFQVDSRSGRGCGTSDNDAAEALNASSKEEAQITREKIPERYLFAFALNSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 3  
 AAB18687  
 ID AAB18687 standard; peptide: 356 AA.  
 AC AAB18687;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Amino acid sequence of a geminivirus replication protein of TGMV.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 354  
 FT /note= "unspecified amino acid"  
 PN WO200054573-A1.  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI: 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant A11 protein with a mutation in the Rb binding region  
PT -  
PS  
PS Disclosure: Page 47-48; 73pp; English.  
XX  
XX  
CC The present sequence represents a geminivirus replication (Rep)  
CC protein, which is also known as A11. A11 binds double-stranded DNA,  
CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the A11 protein are used  
CC to produce transgenic plants. The mutation in A11 is present in a  
CC ribosome binding region, and expression of mutant A11 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant A11  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
XX  
XX  
XX Sequence 356 AA:

Query Match	95.28	Score 340	DB 21	Length 356
Best Local Similarity	95.78	Pred. No. 9.8e-35		
Matches 67	Conservative 0	Mismatches 3	Indels 0	Gaps 0

OY		1	TIVWGEFOVDRSRARGCCGOTSDNDAALNALNASSEKALQIIRKKIPERYLFAAFALNSNL	60
Dd		110	tlvvgefvdgrsrargccgotsnddaaalnassssekallqirrkipekylifgfmlnsnl	169
OY		61	DRFDKTRPEP	70
Dd		170	drlfdtklrpep	179

RESULT	4
AAB18685	
ID	AAB18685 standard; peptide; 70 AA.

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (All) protein.

**KM** Geminivirus: replication protein; Rep protein; All; transgenic plant;  
**KW** ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
OS Tomato golden mosaic virus.

	Key	Location/Qualifiers
FH	Misc-difference	10
FT		/note="wild type residue replaced with Ala"
FT		

PN W0200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR	18-MAR-1999;	99US-0125004.
PR	09-APR-1999;	99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

PI Hanley-Bowdoin L, Orozco BM, Kong L;

DR WPT; 2000-618851/59.

PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
PS  
XX Claim 53: Page 46; 73pp; English.

CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.

50 Sequence 70 AA;

Query Match	93.08;	Score 332;	DB 21;	Length 70;
Best Local Similarity	94.3%;	Pred. No. 1.3e-34;		
Matches	66;	Conservative	0;	Mismatches 4;
			Indels	0;
			Gaps	0;

QY 1 TLVWGEFQVVDGSRARGGCQTSDNDAAEALNASSKEELQIIEIKIPEKYLFAFAALNSNL 60  
|||||  
Db 1 tlvwgefqvavrsargcqtstndaaaalnasskeelqiliekipekylffqfmlnsnl 60

QY	61	DRIEDKTPPEP	70
Db	61	drifdktppep	70

RESULT	5
AAB18688	
ID	AAB18688 standard; peptide: 70 AA

AC AAB18688

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
OS Tomato golden mosaic virus.

	Location/Qualifiers
FH	Key
FT	Misc-difference 19
FT	/note= "wild type residue replaced with Ala"

$$F_T/n$$

PN W0200054573-A1.

PD 21-SEP-2000.

15-MAR-2000; 2000WO-US06759.  
PF  
XX

PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

PI Hanley-Bowdoin L, Orozco BM, Kong L;

DR WPT; 2000-618851/59.

```

XX XX Transgenic plants with increased resistance to geminivirus infection
PT PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX XX
PS PS Disclosure; Page 48; 73pp; English.
XX XX
CC CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX XX
SQ Sequence 70 AA;

Query Match          92.2%; Score 329; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 3.2e-34;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGSRAGCGCOTSDNAAEALNASSKEALQITREKIPKYLPAFAALNSNL 60
   |||||
Db 1 tlwgefdydgdsargcgctsdnaaealnasskealqitrekipekyllfgfmlnsl 60
   |||||

QY 61 DRFDKTPPEP 70
   |||||
Db 61 drfdktpep 70
   |||||

RESULT 6
AAB18692
ID AAB18692 standard; peptide; 70 AA.
XX
AC AAB18692;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
XX OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT MISC-difference 66 /note= "wild type residue replaced with Ala"
FT MISC-difference 69 /note= "wild type residue replaced with Ala"
FT MISC-difference 69 /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
XX 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;

```

```

XX XX WPI; 2000-618851/59.
DR DR
XX XX Transgenic plants with increased resistance to geminivirus infection
PT PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX XX
PS PS Disclosure; Page 50; 73pp; English.
XX XX
CC CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX XX
SQ Sequence 70 AA;

Query Match          91.9%; Score 328; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 4.3e-34;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGSRAGCGCOTSDNAAEALNASSKEALQITREKIPKYLPAFAALNSNL 60
   |||||
Db 1 tlwgefdydgdsargcgctsdnaaealnasskealqitrekipekyllfgfmlnsl 60
   |||||

QY 61 DRFDKTPPEP 70
   |||||
Db 61 drfdktpep 70
   |||||

RESULT 7
AAB18684
ID AAB18684 standard; peptide; 70 AA.
XX
AC AAB18684;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
XX OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT MISC-difference 7 /note= "wild type residue replaced with Ala"
FT MISC-difference 7 /note= "wild type residue replaced with Ala"
FT MISC-difference 8 /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
XX 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.

```

XX PA Hanley-Bowdoin L, Orozco BM, Kong L;  
XX PI WPI: 2000-618851/59.  
XX DR  
XX PT Transgenic plants with increased resistance to geminivirus infection  
XX PT comprise a nucleic acid construct containing a nucleic acid sequence  
XX PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX PS  
XX PS Claim 52: Page 45; 73pp: English.  
XX CC The present sequence represents a mutant peptide, derived from a  
XX CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
XX CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
XX CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
XX CC protein are used to produce transgenic plants. The mutation in AL1 is  
XX CC present in a ribosome binding region, and expression of mutant AL1  
XX CC protein imparts increased resistance to geminivirus infection in the  
XX CC plant. Mutant AL1 proteins are useful for producing plants having  
XX CC increased resistance or reduced sensitivity to a geminivirus such as  
XX CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
XX CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
XX CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
XX CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
XX CC virus, cotton leaf curl virus or beet curly top virus.  
XX SQ Sequence 70 AA:  
SQ  
Query Match 91.3%; Score 326; DB 21: Length 70:  
Best Local Similarity 92.9%; Pred. No. 7.7e-34;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDSARGSCCTSDNDAAEALINASSKEALQIREKIPKYLFAFALINSL 60  
DB 1 TLVWGEAAYGSRGARGCGTSDNDAAEALINASSKEALQIREKIPKYLFGFHNLSN 60  
QY 61 DRIFDKTPEP 70  
DB 61 drifdktp 70  
RESULT 8  
AAB18690  
ID AAB18690 standard; peptide; 70 AA.  
XX AC AAB18690:  
XX DT 22-JAN-2001 (first entry)  
XX DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX KW ribosome binding region; resistance; geminivirus infection.  
XX OS Synthetic.  
XX OS Tomato golden mosaic virus.  
XX FH Key Location/Qualifiers  
XX FT MISC-difference 27 /note= "wild type residue replaced with Ala"  
XX FT MISC-difference 30 /note= "wild type residue replaced with Ala"  
XX FT MISC-difference 30 /note= "wild type residue replaced with Ala"  
XX PN MO200054573-A1.  
XX PD 21-SEP-2000.  
XX PF 15-MAR-2000; 2000MO-US06759.  
XX PR 18-MAR-1999; 9905-0125004.  
XX PR 09-APR-1999; 9905-0289346.  
XX PR

XX PA (UNNC-) UNIV NORTH CAROLINA STATE.  
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX DR WPI: 2000-618851/59.  
XX PT Transgenic plants with increased resistance to geminivirus infection  
XX PT comprise a nucleic acid construct containing a nucleic acid sequence  
XX PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX PS  
XX PS Disclosure: Page 49; 73pp: English.  
XX CC The present sequence represents a mutant peptide, derived from a  
XX CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
XX CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
XX CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
XX CC protein are used to produce transgenic plants. The mutation in AL1 is  
XX CC present in a ribosome binding region, and expression of mutant AL1  
XX CC protein imparts increased resistance to geminivirus infection in the  
XX CC plant. Mutant AL1 proteins are useful for producing plants having  
XX CC increased resistance or reduced sensitivity to a geminivirus such as  
XX CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
XX CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
XX CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
XX CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
XX CC virus, cotton leaf curl virus or beet curly top virus.  
XX SQ Sequence 70 AA:  
SQ  
Query Match 91.3%; Score 326; DB 21: Length 70:  
Best Local Similarity 92.9%; Pred. No. 7.7e-34;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDSARGSCCTSDNDAAEALINASSKEALQIREKIPKYLFAFALINSL 60  
DB 1 TLVWGEFQVDSARGSCCTSDNDAAEALINASSKEALQIREKIPKYLFGFHNLSN 60  
QY 61 DRIFDKTPEP 70  
DB 61 drifdktp 70  
RESULT 9  
AAB18678  
ID AAB18678 standard; peptide; 70 AA.  
XX AC AAB18678:  
XX DT 22-JAN-2001 (first entry)  
XX DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX KW ribosome binding region; resistance; geminivirus infection.  
XX OS Synthetic.  
XX OS Tomato golden mosaic virus.  
XX FH Key Location/Qualifiers  
XX FT MISC-difference 12 /note= "wild type residue replaced with Ala"  
XX FT MISC-difference 13 /note= "wild type residue replaced with Ala"  
XX FT MISC-difference 15 /note= "wild type residue replaced with Ala"  
XX PN MO200054573-A1.  
XX PD 21-SEP-2000.  
XX PR

PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 53; Page 42-43; 73pp; English.  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 Query Match 91.0%; Score 325; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 1e-33;  
 Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVGRSARGCQTSNDAAEALNASSKEEALQIRREKIPKYLFAFALNSNL 60  
 Db 1 tlwgefvgdgrsargcgtsndaaealnaskeaalqirrekipkylfqfhnlnsnl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70  
 RESULT 10  
 ID AAB18686 standard; peptide; 70 AA.  
 AC AAB18686;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 XX Key Location/Qualifiers  
 FT MISC-difference 24 /note= "wild type residue replaced with leu"  
 FT MISC-difference 25 /note= "wild type residue replaced with leu"  
 FT MISC-difference 26 /note= "wild type residue replaced with leu"  
 FT MISC-difference 26 /note= "wild type residue replaced with leu"  
 XX

PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 53; Page 46; 73pp; English.  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 Query Match 91.0%; Score 325; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 1e-33;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVGRSARGCQTSNDAAEALNASSKEEALQIRREKIPKYLFAFALNSNL 60  
 Db 1 tlwgefvgdgrsargcgtsndlllealnaskeaalqirrekipkylfqfhnlnsnl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70  
 RESULT 11  
 ID AAB18689 standard; peptide; 70 AA.  
 AC AAB18689;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 XX Key Location/Qualifiers  
 FT MISC-difference 22 /note= "wild type residue replaced with ala"  
 FT MISC-difference 22 /note= "wild type residue replaced with ala"  
 FT MISC-difference 23

FT		/note= "wild type residue replaced with Ala"
xx		
PN	MO200054573-AI.	
xx		
PD	21-SEP-2000.	
xx		
PF	15-MAR-2000; 2000WO-US06759.	
xx		
PR	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
xx		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
xx		
PI	Hanley-Bowdoin L., Orozco BM, Kong L;	
xx		
DR	WPI: 2000-618851/59.	
xx		
PT	Transgenic plants wltch increased resistance to geminivirus infection	
PT	comprise a nucleic acid construct containing a nucleic acid sequence	
PT	encoding a mutant AI1 protein with a mutation in the Rb binding region	
-		
PS	Disclosure: Page 48-49; 73pp; English.	
xx		
CC	The present sequence represents a mutant peptide, derived from a	
CC	geminivirus replication (Rep) protein, also known as AI1. AI1 binds	
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded	
CC	DNA, and interacts with other viral and host proteins. Mutants of the AI1	
CC	protein are used to produce transgenic plants. The mutation in AI1 is	
CC	present in a ribosome binding region, and expression of mutant AI1	
CC	protein imparts increased resistance to geminivirus infection in the	
CC	plant. Mutant AI1 proteins are useful for producing plants having	
CC	increased resistance or reduced sensitivity to a geminivirus such as	
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl	
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian	
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic	
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper	
CC	virus, cotton leaf curl virus or beet curly top virus.	
xx		
SQ	Sequence 70 AA:	
	Query Match 90.8%; Score 324; DB 21; Length 70;	
	Best Local Similarity 92.9%; Pred No.1.4e-33;	
	Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0	
Oy	1 TLVGEFQVDGSRAGCCCTSDNAAAEALNASSKEBALOIREKIPKYLFAFALNSND 60	
Db	1 IIVVVII 1 IIIII	
	lllvwgefgydgysrargcgqtsaaanaaalnasskcealqlirekipekylfghmnlsl 60	
Oy	61 DRIFDKTPRP 70	
	IIIIIIIIIIII	
Db	61 dRIIdktpep 70	
RESULT 12		
ID AAB18680		
AAB18680 standard; peptide: 70 AA.		
xx		
AC AAB18680;		
xx		
DT 22-JAN-2001 (first entry)		
xx		
DE Mutant peptide derived from amino acids 110-179 of Rep (AI1) protein.		
Geminivirus; Replication protein; Rep protein; AI1; transgenic plant;		
ribosome binding region; resistance; geminivirus infection.		
Synthetic.		
Tomato golden mosaic virus.		
Key Location/Qualifiers		
Misc-difference 42		
FT		

FT		/note= "wild type residue replaced with Ala"
FT	Misc-difference 43	
FT	/note= "wild type residue replaced with Ala"	
FT	Misc-difference 44	
FT	/note= "wild type residue replaced with Ala"	
PN	WO200054573-A1.	
XX		
XX	21-SEP-2000.	
PD		
PF	15-MAR-2000; 2000MO-US06759.	
XX		
XX	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
XX		
P1	Hanley-Bowdoin L, Orozco BM, Kong L;	
XX		
DR	WPI: 2000-618851/59.	
XX		
PT	Transgenic plants with increased resistance to geminivirus infection	
PT	comprise a nucleic acid construct containing a nucleic acid sequence	
PT	encoding a mutant AL1 protein with a mutation in the Rb binding region	
XX		
XX		
PS	Claim 52: Page 43-44; 73pp; English.	
XX		
CC	The present sequence represents a mutant peptide, derived from a	
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds	
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded	
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1	
CC	protein are used to produce transgenic plants. The mutation in AL1 is	
CC	present in a ribosome binding region, and expression of mutant AL1	
CC	protein imparts increased resistance to geminivirus infection in the	
CC	plant. Mutant AL1 proteins are useful for producing plants having	
CC	increased resistance or reduced sensitivity to a geminivirus such as	
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl	
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian	
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic	
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper	
CC	virus, cotton leaf curl virus or beet curly top virus.	
XX		
SO	Sequence 70 AA:	
Query Match	90.2%; Score 322; DB 21; Length 70;	
Best Local Similarity	91.4%; Pred. No. 2.5e-33;	
Matches 64; Conservative	0; Mismatches 6; Indels 0; Gaps 0;	
QY	1 TLWGEFQVNGRKARCCGCTSDAAAEALNMSSKEALQIIRKIPKRYFAPAPALNSL	60
Db	1 TLWGEFQVNGRGRARCGCTSDAAAEALNMSSKEALQIIRKIPKRYFAPAPALNSL	60
QY	61 DRIFFDKTPRP 70	
Db	61 DRIFFDKTPRP 70	
RESULT 13		
AAAB18691		
ID	AAAB18691 standard; peptide; 70 AA.	
XX		
AC	AAAB18691;	
XX		
DT	22-JAN-2001 (first entry)	
XX		
DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.	
XX		
KM	Geminivirus: replication protein; Rep protein; AL1; transgenic plant;	
KW	ribosome binding region; resistance; geminivirus infection.	
XX		
OS	Synthetic.	

```

OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT MISC-difference 34 /note= "wild type residue replaced with Ala"
FT MISC-difference 35 /note= "wild type residue replaced with Ala"
FT MISC-difference 36 /note= "wild type residue replaced with Ala"
FT MISC-difference 36 /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PE 15-MAR-2000; 2000MO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYN-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
PS Disclosure; Page 49; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA:
XX
Query Match 90.2%; Score 322; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 2,5e-33;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
QY 1 TLVWGEFOYDGRSARGCQTSNDAAAEALNASKKEEALQIREKIPKYLFAFALNSNL 60
DB 1 TLVWGEFQVYDGRSARGCQTSNDAAAEALNASKKEEALQIREKIPKYLFIHNL 60
XX
QY 61 DRIFDKTPEP 70
DB 61 DRIFDKTPEP 70
XX
RESULT 14
AAB18681
ID AAB18681 standard; peptide: 70 AA.
XX
AC AAB18681;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX

```

```

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
XX
FH Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT MISC-difference 47 /note= "wild type residue replaced with Ala"
FT MISC-difference 48 /note= "wild type residue replaced with Ala"
FT MISC-difference 49 /note= "wild type residue replaced with Ala"
FT MISC-difference 49 /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PE 15-MAR-2000; 2000MO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYN-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
PS Claim 52; Page 44; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA:
XX
Query Match 89.4%; Score 319; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 6e-33;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
QY 1 TLVWGEFOYDGRSARGCQTSNDAAAEALNASKKEEALQIREKIPKYLFAFALNSNL 60
DB 1 TLVWGEFQVYDGRSARGCQTSNDAAAEALNASKKEEALQIREKIPKYLFIHNL 60
XX
QY 61 DRIFDKTPEP 70
DB 61 DRIFDKTPEP 70
XX
RESULT 15
AAB18683
ID AAB18683 standard; peptide: 70 AA.
XX
AC AAB18683;
XX

```



Job time: 527 sec

```

DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key location/Qualifiers
FT Misc-difference 59 /note= "wild type residue replaced with Ala"
FT Misc-difference 61 /note= "wild type residue replaced with Ala"
FT Misc-difference 62 /note= "wild type residue replaced with Ala"
FT Misc-difference 62 /note= "wild type residue replaced with Ala"
XX
PN W0200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX 09-APR-1999; 99US-0289346.
XX
XX (UYN-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DK WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rd binding region
PT
XX
XX Claim 53; Page 45; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

```

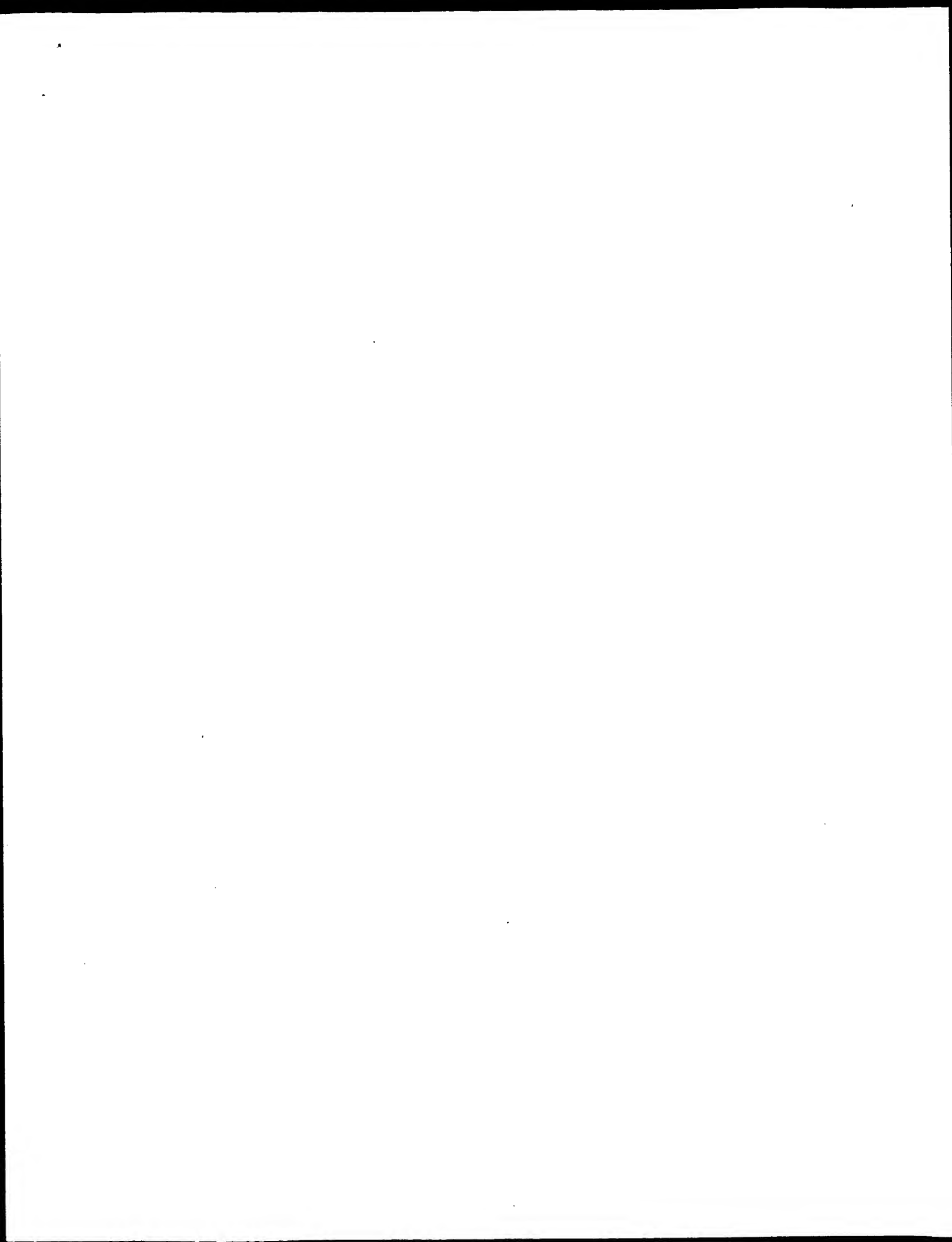
Query Match 89.1%; Score 318; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 8e-33;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 TLVWGEQVYGRSARGGCCOTSNDAAEALNASSKEEALQITREKIPKYLFAFAALNSN. 60
   |||||||
DB 1 tlwgeqyvqrsargcgctsndaaealnasskeeaqlitrekipekylfqtfnlnsal 60
   |||||||
QY 61 DRIFDKTPEP 70
   |||||||
DB 61 aaifdktppep 70

```

Search completed: October 10, 2002, 01:47:18



Page 1

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds  
(without alignments)

(without alignments)  
102.295 million cell updates/sec

Title: US-09-289-346A-7

Perfect score: 359  
Sequence: I TLVWGEFYVDGRSARGGQT.....FQFHNLNSALAIPTKTPP 70

Scoring table:	BLOSUM62	Canext	0.5
	Canext	10.0	

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

1:  sp.archaea:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp.human:*
5:  sp.invertebrate:*
6:  sp.mammal:*
7:  sp.mhc:*
8:  sp.organelle:*
9:  sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vetlebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	293	81.6	351	12	Q91R10	Q91I10	tomato seve
2	291	81.1	352	12	Q9E000	Q9E000	tomato rogo
3	274	76.3	226	12	O09727	O09727	leonorus mo
4	274	76.3	226	12	Q9WHF6	Q9WHF6	tomato mild
5	272	75.8	361	12	O67574	O67574	bean golden
6	270	75.2	225	12	Q9QDB1	Q9QDB1	cowpea gold
7	270	75.2	314	12	Q9ELT8	Q9ELT8	sweet potat
8	270	75.2	364	12	Q9Q555	Q9Q555	sweet potat
9	267	74.4	185	12	Q98693	Q98693	sids golden
10	259	72.1	149	12	P88975	P88975	macropiliu
11	259	72.1	233	12	Q9YL44	Q9YL44	macropiliu
12	258	71.9	234	12	Q93180	Q93180	geminiivird
13	252	70.2	190	12	Q92089	Q92089	tomacco lea
14	252	70.2	190	12	Q92084	Q92084	tomacco lea
15	245	69.4	190	12	Q9W827	Q9W827	tomacco lea
16	245	69.4	208	12	Q920C4	Q920C4	tomacco lea

17	249	69.4	353	12	072692	076692	beet curly
18	249	69.4	354	12	065438	065438	beet curly
19	249	69.4	359	12	091m88	091m88	tabacco lea
20	247	68.8	208	12	0920C0	0920C0	tabacco lea
21	247	68.8	208	12	0920B8	0920B8	tabacco lea
22	246	68.5	203	12	092083	0920b6	tabacco lea
23	245	68.2	208	12	0920B6	0920b6	tabacco lea
24	245	68.2	358	12	065418	065418	beet curly
25	244	68.0	363	12	073577	073577	cotton leaf
26	239	66.6	359	12	091B86	091B86	ageratum ye
27	239	66.6	360	12	090DX0	090DX0	ageratum ye
28	238	66.3	190	12	092086	092086	tabacco lea
29	238	66.3	362	12	0911m5	09115	macroptiliu
30	238	66.3	348	12	056816	056816	chayote mos
31	237	66.0	208	12	0920A0	0920A0	tabacco lea
32	237	66.0	363	12	091M48	091M48	okra enatio
33	237	66.0	362	12	072705	072705	cotton leaf
34	236	66.0	363	12	072719	072719	cotton leaf
35	236	65.7	208	12	0920C6	0920C6	tabacco lea
36	235	65.5	349	12	088888	088888	tomato pseu
37	234	65.2	359	12	092v44	092v44	tomato yell
38	234	65.2	359	12	092v22	092v22	tomato yell
39	234	65.2	359	12	092vX7	092vX7	tomato yell
40	233	64.9	359	12	088942	088942	tomato yell
41	233	64.6	190	12	0920A7	0920a7	tabacco lea
42	231	64.3	307	12	091E77	091E77	cotton leaf
43	231	64.3	361	12	072723	072723	cotton leaf
44	230	64.1	231	12	096620	096620	african tom
45							

## ALIGNMENTS

RESULT	1
Q91R10	
ID	Q91R10
AC	Q91R10;
DT	01-DEC-2001 (TREMBLrel, 19, Created)
DT	01-DEC-2001 (TREMBLrel, 19, Last sequence update)
DT	01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DE	REP PROTEIN.
GN	AC1.
OS	Tomato severe rugose virus.
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX	NCBI_TaxID=158463;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	SPPAIN-MTNAS GERAIS;
RA	Requena W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT	"The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT	begomovirus, Tomato severe rugose virus, in Brazil."
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AY029750; AAK50357.1; -
SO	SEQUENCE 351 AA: 40122 MW: 87P93744F873B6CF CR664;

	Query Match	81.6%	Score 293;	DB 12;	Length 351;
	Best Local Similarity	78.6%;	Pred. No. 2.9e-25;		
	Matches 55; Conservative	7;	Mismatches 8;	Indels 0;	Gaps 0;
OY	1 TLVGEFQVDGSGARCGCOTSDNDAAEALNAASKEEALQIREKIPEKYLPOFHNLNSAL	60			
	:				
Db	111 TLEWGFQIDGSGARCGCOTANDAAEAALNAESKVALQIIREKLPERKLPFOHNLNSVL	170			
OY	61 AAIFDKTPEP	70			
	:				
Db	171 DRIFARAPEP	180			
RESULT	2				
Q9BE00					
ID Q9BE000	PRELIMINARY;	PRT:	352	AA.	

AC Q9E000;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN.  
 GN ACI  
 OS Tomato rugose mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_Taxid=134599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fernandes J.J., Fontes F.P.B., Brommonschenkel S.H., Carvalho M.G.,  
 RA Zambolim E.M., Zerbini F.M.;  
 RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
 RT (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,  
 RT Minas Gerais, Brazil."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF291705; AAG15546.1; -  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 SQ SEQUENCE 352 AA; 40012 MW; 47CD55838E24D613 CRC64;

Query Match 81.1%; Score 291; DB 12; Length 352;  
 Best Local Similarity 78.6%; Pred. No. 4.9e-25;  
 Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCQTSNDAAAEALNASSKEEALQITREKIPKEYLFOFHNLNSAL 60  
 Db 111 TIEMGEFQIDGRSARGCGQTANDAAAEALNAPSKDIALQITREKMPERKLFQFHNLNSL 170  
 QY 61 AAIEDKTPPEP 70  
 Db 171 DRIFAKAPEP 180

RESULT 3  
 009727  
 ID 009727 PRELIMINARY; PRT; 226 AA.  
 AC 009727;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonurus mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_Taxid=58177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-LEUV- BRAZIL 1;  
 RA Faria J.C., Maxwell D.P.;  
 RT "Variability in geminivirus associated with Phaseolus vulgaris in  
 RT Brazil."  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U92532; AAB51157.1; -  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 226  
 PT 226  
 SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 76.3%; Score 274; DB 12; Length 226;  
 Best Local Similarity 77.1%; Pred. No. 2.5e-23;  
 Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCQTSNDAAAEALNASSKEEALQITREKIPKEYLFOFHNLNSAL 60  
 Db 111 TIEMGEFQVDRSARGCGQTANDAAAEALNAPDKTALQITREKIPKEYLFOFHNLNSL 170

QY 61 AAIEDKTPPEP 70  
 Db 171 DRIFAKAPEP 180

RESULT 4  
 09WHF6  
 ID 09WHF6 PRELIMINARY; PRT; 226 AA.  
 AC 09WHF6;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Tomato mild mottle geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
 OX NCBI\_Taxid=92943;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-HN96-H5KW;  
 RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
 RA Maxwell D.P.;  
 RT "Molecular characterization and DNA-based detection methods for  
 RT vegetable-infecting geminiviruses in Central America."  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF131071; AAD33471.1; -  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 DR NON\_TER 226  
 FT 226  
 SQ SEQUENCE 226 AA; 25941 MW; 2EA1116712871A23 CRC64;

Query Match 76.3%; Score 274; DB 12; Length 226;  
 Best Local Similarity 72.9%; Pred. No. 2.5e-23;  
 Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCQTSNDAAAEALNASSKEEALQITREKIPKEYLFOFHNLNSAL 60  
 Db 111 TIEMGEFQIDGRSARGCGQTANDAAAEALNASSKEEALQITREKIPKEYLFOFHNLNSL 170  
 QY 61 AAIEDKTPPEP 70  
 Db 171 DRIFAKAPEP 180

RESULT 5  
 067574  
 ID 067574 PRELIMINARY; PRT; 361 AA.  
 AC 067574;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN ALI.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_Taxid=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 RA Maxwell D.P., Russell D.R.;  
 RT "Cloning of the complete DNA genomes of four bean-infecting  
 RT geminiviruses and determining their infectivity by electric discharge  
 RT particle acceleration."  
 RL Phytopathology 81:980-985(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 RA Morales F.J., Maxwell D.P.;  
 RT "Differentiation of bean-infecting geminiviruses by nucleic acid

RT hybridization probes and aspects of bean golden mosaic in Brazil.,"  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.:  
 RT "Genetic diversity in geminiviruses causing bean golden mosaic  
 RT disease: The nucleotide sequence of the infectious cloned DNA  
 RT components of a Brazilian isolate of bean golden mosaic geminivirus."  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M88686; AAA6312.1; -  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemin1\_AL1; 1.  
 SO SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 75.8%; Score 272; DB 12; Length 361;  
 Best Local Similarity 76.1%; Pred. No. 7.3e-23;  
 Matches 51; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

OY 4 WGEFVODRSARGCGCTNDAAAEALNASSKEEALQITREKIPKYLFOFHNLNSALAI 63  
 Db 113 WGEFVODRSARGCGCTNDAAAEALNASSKEEAMQIIKELPKFLFYHNLSSNLDRI 172  
 OY 64 FDKTRPEP 70  
 Db 173 FTKRDP 179

RESULT 6  
 O9QDB1  
 ID O9QDB1 PRELIMINARY: PRT: 225 AA.  
 AC O9QDB1.  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=69263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CGMV-BR;  
 RA Faria J.C.:  
 RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from  
 RT Brazil."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF188708; AAF06318.1; -  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemin1\_AL1; 1.  
 FT NON\_TER 225 225  
 SO SEQUENCE 225 AA; 25766 MW; 1089CH6BD8D15B5D CRC64;

Query Match 75.2%; Score 270; DB 12; Length 225;  
 Best Local Similarity 76.1%; Pred. No. 7.2e-23;  
 Matches 51; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 4 WGEFVODRSARGCGCTNDAAAEALNASSKEEALQITREKIPKYLFOFHNLNSALAI 63  
 Db 113 WGEFVODRSARGCGCTNDAAAEALNASSKEEAMQIIKELPKFLFYHNLSSNLDRI 172  
 OY 64 FDKTRPEP 70  
 Db 173 FKKRPEP 179

RESULT 7  
 O9ELT8

ID O9ELT8 PRELIMINARY: PRT: 314 AA.  
 AC O9ELT8;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.:  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 RT States."  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A. to the EMBL/GenBank/DBJ databases.  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF288227; AAG01006.1; -  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemin1\_AL1; 1.  
 SO SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 75.2%; Score 270; DB 12; Length 314;  
 Best Local Similarity 79.4%; Pred. No. 1e-22;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 1 TIVWGEFVODRSARGCGCTNDAAAEALNASSKEEALQITREKIPKYLFOFHNLNSAL 60  
 Db 110 TIVWGEFVODRSARGCGCTNDAAAEALNASSKEEALQITREKIPKYLFOFHNLNSAL 169  
 OY 61 AALFDKTP 68  
 Db 170 DRIFSPPP 177

RESULT 8  
 O9QSS5  
 ID O9QSS5 PRELIMINARY: PRT: 364 AA.  
 AC O9QSS5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.:  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 RT States."  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.:  
 RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl  
 RT virus : nucleotide sequence and phylogenetic relationships."  
 RL Online Publication.  
 DR EMBL: AF104036; AAD47173.1; -  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemin1\_AL1; 1.  
 SO SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 75.2%; Score 270; DB 12; Length 364;  
 Best Local Similarity 79.4%; Pred. No. 1.2e-22;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 60  
 110 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 169

DB 61 AAFDKTPP 68  
 170 DRIFSKPP 177

RESULT 9  
 098693 PRELIMINARY; PRT; 185 AA.

AC 098693;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN ACl.  
 OS sida golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51034;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
 RT Genetic Diversity among geminiviruses associated with the weed  
 RT species sida spp, Macroplidium lathyroides, and Wissadula amplissima  
 RT from Jamaica.";  
 RL Plant Dis. 81:1251-1258(1997).  
 DR EMBL: U67926; AAB97865.1; -;  
 DR Interpro: IPR001191; GeminI\_AL1.  
 DR Pfam: PF00799; GeminI\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; GeminI\_AL1; 1.  
 FT NON\_TER 1 185  
 FT SEQUENCE 185 AA; 20975 MW; 3913850A025A5EF1 CRC64;

Query Match 74.4%; Score 267; DB 12; Length 185;  
 Best Local Similarity 70.0%; Pred. No. 1.3e-22;  
 Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 60  
 89 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 148

QY 61 AAFDKTPP 70  
 149 DRIFSKPP 158

RESULT 10  
 P88975 PRELIMINARY; PRT; 149 AA.

AC P88975;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN ACl.  
 OS Macroplidium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICAN;  
 RA Roye M.E.;  
 RL Thesis (1996), Biochemistry, University of the West Indies, Jamaica.

DR EMBL: U75278; AAB36919.1; -;  
 DR Interpro: IPR001191; GeminI\_AL1.  
 DR Pfam: PF00799; GeminI\_AL1;  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; GeminI\_AL1; 1.  
 FT NON\_TER 1 149  
 FT SEQUENCE 149 AA; 16785 MW; E4CF5ED4C9CD508 CRC64;

Query Match 72.1%; Score 259; DB 12; Length 149;  
 Best Local Similarity 68.6%; Pred. No. 8e-22;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 60  
 52 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 111

QY 61 AAFDKTPP 70  
 112 DRIFSKPP 121

RESULT 11  
 09YL4 PRELIMINARY; PRT; 233 AA.

AC 09YL4;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Macroplidium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.;  
 RT "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses  
 RT from Jamaica.";  
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E., McLaughlin W.A., Maxwell D.P.;  
 RT "Molecular characterization of two distinct geminiviruses infecting M.  
 RT lathyroides from Jamaica.";  
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF098940; AAD17850.1; -;  
 DR Interpro: IPR001191; GeminI\_AL1.  
 DR Pfam: PF00799; GeminI\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; GeminI\_AL1; 1.  
 FT NON\_TER 233  
 FT SEQUENCE 233 AA; 26356 MW; AA490AF4D216A02 CRC64;

Query Match 72.1%; Score 259; DB 12; Length 233;  
 Best Local Similarity 68.6%; Pred. No. 1.3e-21;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 60  
 110 TLWGEFQVDSRARGCGCOTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLNSAL 169

QY 61 AAFDKTPP 70  
 170 DRIFSKPP 179

RESULT 12  
 039180

ID 039180 PRELIMINARY; PRT: 234 AA.  
 AC 039180;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE AC1 (FRAGMENT).  
 OS Geminiviridae.  
 OC Viruses; ssDNA viruses.  
 OX NCBI\_TaxID=10811;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TOMATO STRAIN.  
 RA Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;  
 RT "Partial Characterization of Two Whitefly-Transmitted Geminiviruses  
 Infecting Tomatoes in Venezuela."  
 RL Plant Dis. 81:312-312(1997).  
 DR EMBL: AF026553; AAB82605.1; -  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Geminl\_AL1: 1.  
 FT NON\_TER 234  
 FT 234  
 SQ SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;

Query Match 71.9%; Score 258; DB 12; Length 234;  
 Best Local Similarity 68.6%; Pred. No. 1.7e-21;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGCGCOTSDNDAAEALNASSKEALQIRREKIPKYLFOFHNSAL 60  
 DB 110 TLEWGFQIDGSRAGCGQOTANDAAAEALNASTGTEAKKIKRKLPEKFLFOYHNLSL 169  
 61 AAFDKTPEP 70  
 170 DRIFMKAPNP 179

RESULT 13  
 Q92089 PRELIMINARY; PRT: 190 AA.  
 ID 092089  
 AC 092089;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL  
 AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YOKOHAMA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1; -  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Geminl\_AL1: 1.  
 FT NON\_TER 190  
 FT 190  
 SQ SEQUENCE 190 AA; 21432 MW; AAC093D1D1610FAD CRC64;

Query Match 70.2%; Score 252; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 6.6e-21;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 QY 1 TLVWGEFQVDSARGCGCOTSDNDAAEALNASSKEALQIRREKIPKYLFOFHNSAL 60  
 110 TLEWGFQIDGSRAGCGQOTANDAAAEALNASTGTEAKKIKRKLPEKFLFOYHNLSL 169

DB 85 TLEWGFQIDGSRAGCGQONADACAEALNASSKEALATIRREKLPKDFIYOHNSNL 144  
 QY 61 -----AAFDKTPPE 69  
 DB 145 DRIFAPPELVFCPTASSFDVPE 169

RESULT 14  
 ID 092084 PRELIMINARY; PRT: 190 AA.  
 AC 092084;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL  
 AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YOKOHAMA5;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001318; BAA34039.1; -  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Geminl\_AL1: 1.  
 FT NON\_TER 190  
 FT 190  
 SQ SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

Query Match 70.2%; Score 252; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 6.6e-21;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDSARGCGCOTSDNDAAEALNASSKEALQIRREKIPKYLFOFHNSAL 60  
 DB 85 TLEWGFQIDGSRAGCGQONADACAEALNASSKEALATIRREKLPKDFIYOHNSNL 144  
 QY 61 -----AAFDKTPPE 69  
 DB 145 DRIFAPPELVFCPTASSFDVPE 169

RESULT 15  
 ID 09W827 PRELIMINARY; PRT: 190 AA.  
 AC 09W827;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORAL-1, PARTIAL AND  
 COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GORAL;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001303; BAA34010.1; -  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR Prodom: PD000736; Geminl\_AL1: 1.

FT NON\_TER 1 1  
 FT NON\_TER 190 190  
 SQ SEQUENCE 190 AA: 21444 MW: 93C3742A8FBD7EB CRC64:

Query Match 69.4%; Score 249; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 1.5e-20;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVWGEFOVDGRSARGCCOTSNDAAAEALNASSKEEALQIIRKTIPEKYLFOPHNLNSAL 60  
 |||||  
 Db 85 TLEWGTFOVDGRSARGCCONANDACAEALNASKAEALSIIREKLPKDFIFOYHNLNSNL 144  
 QY 61 AAI-----FDKTPE 69  
 |||||  
 Db 145 DRIFAPLELVFCPPSSNSFDQVPE 169

Search completed: October 10, 2002, 02:02:00  
 Job time: 1093 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:52 ; Search time 34.62 Seconds  
(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346a-7

Perfect score: 359  
Sequence: 1 TLWGEFQVQGRSARGGCGT.....PQFHNLSALAIPEKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	342	95.3	352	1 VAL1_TGMV	P03567 tomato gold
2	242	67.4	361	1 VAL1_PYMV	P27258 potato yell
3	229	63.8	358	1 VAL1_CLVK	P14982 cassava lat
4	229	63.8	358	1 VAL1_CLVN	P14972 cassava lat
5	222	61.8	362	1 VAL1_TYLCV	P36279 tomato yell
6	216	60.2	349	1 VAL1_PHYV	P06923 pepper huas
7	215	59.9	359	1 VAL1_TYLCU	P38609 tomato yell
8	212	59.1	358	1 VAL1_BCTV	P14991 beet curly
9	211	58.8	353	1 VAL1_BGVV	P01575 bean golden
10	211	58.8	359	1 VAL1_TYLCV	P27260 tomato yell
11	210	58.5	355	1 VAL1_ABYVM	P21947 abutilon mo
12	207	57.7	361	1 VAL1_TMYV	P06657 tomato molt
13	193	53.8	357	1 VAL1_TYLCV	P27259 tomato yell
14	107	29.8	347	1 VAL1_SLVC	P29048 squash leaf
15	68	18.9	447	1 TBL2_HUMAN	Q94433 homo sapien
16	63.5	17.7	1610	1 CCAD_MESAU	Q99244 mesocricetu
17	63.5	17.7	2161	1 CCAD_HUMAN	P07732 ratu
18	63.5	17.7	2203	1 CCAD_RAT	P27732 ratu
19	61	17.0	1483	1 CYP1_YEAST	P12351 saccharomyc
20	59.5	16.6	224	1 SP2R_BACSU	P39151 bacillus su
21	59.5	16.6	4466	1 DYHC_TRIGR	P23098 tritripentes
22	59	16.4	387	1 VE2_HPV1	P27552 human papil
23	59	16.4	387	1 V4PE_RHISN	P55615 rhizobium s
24	58.5	16.3	136	1 V452_CAEEL	O62250 caenorhabdi
25	58.5	16.3	299	1 Y175_HELPJ	Q92mg7 helioabacte
26	57.5	16.0	367	1 LHX4_MOUSE	P33776 mus musculu
27	57.5	16.0	1852	1 CCAS_CYPCA	P23316 cyprinus ca
28	57	15.9	247	1 YCP4_YEAST	P25349 saccharomyc
29	57	15.9	256	1 YKJ9_YEAST	P34247 saccharomyc
30	56.5	15.7	511	1 HUTH_VIRCH	Q9ksq4 vibrio chol
31	56.5	15.7	2190	1 CCAD_CHICK	O73700 gallus gall
32	56	15.6	98	1 V4TF_METJA	O57671 methanococc
33	56	15.6	295	1 VAL1_CSMV	P18921 chloris str

34	56	15.6	295	1 VAL1_TYDVA	P31617 tobacco yel
35	56	15.6	492	1 M0T3_MOUSE	O35308 mus musculu
36	56	15.6	492	1 M0T3_RAT	O70461 rattus norv
37	56	15.6	836	1 GCSR_HUMAN	O99062 homo sapien
38	55.5	15.5	266	1 ETPB_MYCLE	O33093 mycobacteri
39	55.5	15.5	266	1 ETPB_MYCTU	O53276 mycobacteri
40	55.5	15.5	428	1 GBAL_CANAL	P28868 candida alb
41	55.5	15.5	887	1 MTP_BOVIN	P55156 bos taurus
42	55.5	15.5	1066	1 D295_HUMAN	O9ulj3 homo sapien
43	55.5	15.5	4466	1 DYHC_ANTCR	P39057 anthracidari
44	55	15.3	354	1 CARA_METJA	O58425 methanococc
45	55	15.3	481	1 KNIR_DROVI	O24753 drosophila

## ALIGNMENTS

RESULT	ID	VAL1_TGMV	STANDARD	PRT	352 AA
AC	P03567	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	01-JUN-1994 (Rel. 29, Last annotation update)				
DE	ALI protein.				
GN	AC1.				
OS	Tomato golden mosaic virus (TGMV).				
OC	Viruses: ssDNA viruses; Geminiviridae; Begomovirus.				
ON	NCBI_TaxId=10831;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;				
RT	"Complete nucleotide sequence of the infectious cloned DNA components of tomato golden mosaic virus: potential coding regions and regulatory sequences."				
RL	EMBO J. 3:2197-2205 (1984).				
CC	-I- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	EMBL: K02029; -; NOT_ANNOTATED_CDS.				
DR	PIR: A04170; Q0CVL1.				
DR	InterPro: IPR001191; Geminivir_1.				
DR	Pfam: PF00799; Geminivir_1.				
DR	PRINTS: PR00227; GEMINIVIR_1.				
DR	PRODOM: PD000736; Geminivir_1.				
KW	ATP-binding.				
FT	NP-BIND 223				
SO	SEQUENCE 352 AA; 40332 MW; C33C938E9644BA4 CRC64;				
Query Match	95.3%; Score 342; DB 1; Length 352;				
Best Local Similarity	95.7%; Pred. No. 1.7e-31;				
Matches	67; Conservative 0; Mismatches 3; Indels 0;				
QY	1 TLWGEFQVQGRSARGGCGTSDAAAEALNASSKEEALQIRREKPEKYLQFQHNLSAL 60				
DB	111 TLWGEFQVQGRSARGGCGTSDAAAEALNASSKEEALQIRREKPEKYLQFQHNLSNL 170				
QY	61 AAIFDKTPPEP 70				
DB	171 DRIFDKTPPEP 180				
RESULT	2				
VAL1_PYMV	STANDARD;				
ID	VAL1_PYMV	PRT;	361 AA.		

AC P27258;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE AL1 protein.  
 OS Potato yellow mosaic virus (isolate Venezuela).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10828;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91311403; PubMed=1856690;  
 RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;  
 RT "The nucleotide sequence of the infectious cloned DNA components of  
 RL J. Gen. Virol. 72:1515-1520(1991).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMBL; D00940; BAA00782.1; -  
 DR PIR; J00364; QOCVPT.  
 DR InterPro; IPR001191; Geminl\_AL1.  
 DR Pfam; PF00799; Geminl\_AL1.1.  
 DR PRINTS; PR00227; GEMCOATAL1.  
 DR ProDom; PD000736; Geminl\_AL1.1.  
 DR ATP-binding.  
 FT NP\_BIND 222 229 ATP (POTENTIAL).  
 SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match 67.4%; Score 242; DB 1; Length 361;  
 Best Local Similarity 66.7%; Pred. No. 3.8e-20;  
 Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGSRAGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSAL 60  
 DB 110 TTEMWGFQIDGSRAGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSAL 169

OY 61 AATFDKTPPE 69  
 DB 170 DRIFMKAPE 178

RESULT 3  
 VAL1\_CLVKN STANDARD; PRT; 358 AA.  
 AC P14982;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein (40.4 kDa protein).  
 GN AC1.  
 OS Cassava latent virus (strain West Kenyan 844).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Gay M.R.;  
 RT "Nucleotide sequence of cassava latent virus DNA."  
 RL Nature 301:260-262(1983)  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMBL; J02057; -; NOT\_ANNOTATED\_CDS.  
 DR InterPro; IPR001191; Geminl\_AL1.  
 DR Pfam; PF00799; Geminl\_AL1.1.  
 DR PRINTS; PR00227; GEMCOATAL1.  
 DR ProDom; PD000736; Geminl\_AL1.1.  
 DR ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 63.8%; Score 229; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 1.1e-18;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGSRAGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSAL 60  
 DB 109 TTEMWGFQIDGSRAGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSAL 168

OY 61 AATFDKTPPE 70  
 DB 169 DRIFKEPPAP 178

RESULT 4  
 VAL1\_CLVKN STANDARD; PRT; 358 AA.  
 AC P14972;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein (40.4 kDa protein).  
 GN AC1.  
 OS Cassava latent virus (strain Nigerian).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10819;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90174930; PubMed=2308831;  
 RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;  
 RT "Nucleotide sequence of the infectious cloned DNA components of  
 RL African cassava mosaic virus (Nigerian strain)."  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMBL; X17095; CAA34953.1; -  
 DR PIR; S07594; S07594.  
 DR InterPro; IPR001191; Geminl\_AL1.  
 DR Pfam; PF00799; Geminl\_AL1.1.  
 DR PRINTS; PR00227; GEMCOATAL1.  
 DR ProDom; PD000736; Geminl\_AL1.1.  
 DR ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA; 40435 MW; 1DB16B0CB2D5E2C CRC64;

Query Match 63.8%; Score 229; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 1.1e-18;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGSRAGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSAL 60  
 DB 109 TTEMWGFQIDGSRAGCGCOTSDNDAAEALNASSKEEALQIREKIPKYLQFPHNLNSAL 168

Qy 61 AAIPEKTPPEP 70  
Db 169 DRIFQPEPPAP 178

## RESULT 5

VALI\_TYLCU STANDARD: PRT: 362 AA.  
AC P36279;  
DR 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
CL AL1 protein (C1 protein).  
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=36447;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93139778; PubMed=8423446;  
RA Dry I.B., Rigden J.E., Krahe L.R., Mullineaux P.M., Rezaian M.A.;  
RT "Nucleotide sequence and genome organization of tomato leaf curl  
geminivirus.";  
RL J. Gen. Virol. 74:147-151(1993).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
DR PIR: J01887; J01887.  
DR InterPro: IPR001191; Geminl\_AL1.  
DR Pfam: PF00799; Geminl\_AL1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR Prodom: PD000736; Geminl\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 221 228 ATP (POTENTIAL).  
SQ SEQUENCE 362 AA: 41197 MW: 343E7184B4704098 CRC64;

Query Match 61.8%; Score 222; DB 1; Length 362;  
Best Local Similarity 52.9%; Pred. No. 7, 1e-18;  
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

Qy 1 TLWGEFQVDRSARGGCGTSDNAAEALNMSKKEALQIREKIPKYLFOFHNLSAL 60  
Db 110 TLWGEFQVDRSARGGCGTSDNAAEALNMSKKEALQIREKIPKYLFOFHNLSAL 169  
Qy 61 AAI-----FKTPPE 69  
Db 170 DRIFQPEPPAP 178

## RESULT 6

VALI\_PHVU STANDARD: PRT: 349 AA.  
AC Q06923;  
DR 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DE 01-JUN-1994 (Rel. 29, Last annotation update)  
CL AL1 protein.  
OS Pepper huasteco virus (PHV).  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=28349;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94015007; PubMed=8409944;  
RA Torres-Pacheco I., Garzon-Fernando J.A., Herrera-Estrella L.,  
RT Rivera-Bustamante R.F.;  
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and  
comparison with bipartite geminiviruses.";  
RL J. Gen. Virol. 74:2225-2231(1993).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X70418; CAA49856.1; -  
DR PIR: S31875; S31875.  
DR PIR: J02300; J02300.  
DR InterPro: IPR001191; Geminl\_AL1.  
DR Pfam: PF00799; Geminl\_AL1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR Prodom: PD000736; Geminl\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 221 228 ATP (BY SIMILARITY).  
SQ SEQUENCE 349 AA: 39722 MW: D5FAE76CD56370F4 CRC64;

Query Match 60.2%; Score 216; DB 1; Length 349;  
Best Local Similarity 58.6%; Pred. No. 3, 3e-17;  
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TLWGEFQVDRSARGGCGTSDNAAEALNMSKKEALQIREKIPKYLFOFHNLSAL 60  
Db 110 TLWGEFQVDRSARGGCGTSDNAAEALNMSKKEALQIREKIPKYLFOFHNLSAL 169  
Qy 61 AAIPEKTPPEP 70  
Db 170 NRIFQTPPEP 179

## RESULT 7

VALI\_TYLCU STANDARD: PRT: 359 AA.  
AC P38609;  
DR 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DE 01-OCT-1994 (Rel. 30, Last annotation update)  
CL AL1 protein (C1 protein).  
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=37139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94256836; PubMed=8198442;  
RA Nodis E., Hidalgo E., Accotto G., Moriones E.;  
RT "High similarity among the tomato yellow leaf curl virus isolates  
from the west Mediterranean basin: the nucleotide sequence of an  
infectious clone from Spain.";  
RL Arch. Virol. 135:165-170(1994).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z25751; CAA81026.1; -  
DR PIR: S39211; S39211.  
DR InterPro: IPR001191; Geminl\_AL1.  
DR Pfam: PF00799; Geminl\_AL1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR Prodom: PD000736; Geminl\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 221 228 ATP (POTENTIAL).  
SQ SEQUENCE 359 AA: 41065 MW: 2D170A51EFD0A3BC CRC64;

Query Match 59.9%; Score 215; DB 1; Length 359;

Best Local Similarity 58.0%; Pred. No. 4,4e-17;  
Matches 40; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVMGEFOVDGSRAGCGCOTSDNDAAEALNASSKEPALQIREKIPKYLFOFHNLNSAL 61  
DB 111 LVMGEFOVDGSRAGCGCOTSDNDAAEALNASSKEPALQIREKIPKYLFOFHNLNSAL 170

QY 62 AIFDKTPEP 70  
DB 171 RVFGVPPAP 179

RESULT 8  
VAL1\_BCTV STANDARD; PRT; 358 AA.  
AC P14991;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE AL1 protein (40.8 kDa protein).  
OS Beet curly top virus (BCTV).  
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
OX NCBI\_TaxID=10840;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;  
RT "The nucleotide sequence of an infectious clone of the geminivirus  
RT beet curly top virus".  
RL EMBL J. 5:1761-1767(1986).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X04144; -; NOT\_ANNOTATED\_CDS.  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATAL1.  
DR ProDom: PD000736; Gemini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 222 229 ATP (POTENTIAL).  
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C33 CRC64;

Query Match 59.1%; Score 212; DB 1; Length 358;  
Best Local Similarity 55.7%; Pred. No. 9.5e-17;  
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFOVDGSRAGCGCOTSDNDAAEALNASSKEPALQIREKIPKYLFOFHNLNSAL 60  
DB 110 TLVWGEFOVDGSRAGCGCOTSDNDAAEALNASSKEPALQIREKIPKYLFOFHNLNSAL 169

QY 61 AIFDKTPEP 70  
DB 170 QKIFORPPDP 179

RESULT 9  
VAL1\_BGMV STANDARD; PRT; 353 AA.  
AC P05175;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE AL1 protein (40.2 kDa protein).  
OS Bean golden mosaic virus.  
GN AC1.  
DR Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
JC

OX NCBI\_TaxID=10839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Howarth A.J., Caton J., Bossett M., Goodman R.M.;  
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene  
RT regulation in geminiviruses".  
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M10070; AAA46318.1; -  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATAL1.  
DR ProDom: PD000736; Gemini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 222 229 ATP (POTENTIAL).  
SQ SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match 58.8%; Score 211; DB 1; Length 353;  
Best Local Similarity 58.6%; Pred. No. 1.2e-16;  
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFOVDGSRAGCGCOTSDNDAAEALNASSKEPALQIREKIPKYLFOFHNLNSAL 60  
DB 110 TLVWGEFOVDGSRAGCGCOTSDNDAAEALNASSKEPALQIREKIPKYLFOFHNLNSAL 169

QY 61 AIFDKTPEP 70  
DB 170 ERIFKVPDP 179

RESULT 10  
VAL1\_TYLCM STANDARD; PRT; 359 AA.  
AC P27260;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 protein (C1 protein).  
GN C1.  
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=9210760; PubMed=1840676;  
RA Kheyr-Pour A., Bendahmane M., Matzelt V., Accotto G.P., Crespi S.,  
RA Groenborn B.;  
RT "Tomato yellow leaf curl virus from Sardinia is a  
RT whitefly-transmitted monopartite geminivirus".  
RL Nucleic Acids Res. 19:6763-6769(1991).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X61153; CAA43466.1; -  
DR PIR: S22593; S22593  
DR InterPro: IPR001191; Gemini\_AL1.

DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR Prodom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 359 AA; 40733 MW; 9717B4A07C93EFA7 CRC64;

Query Match 58.8%; Score 211; DB 1; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 1,2e-16;  
 Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

OY 2 LVMGEFQVDGSRAGCGCOTSDAAAEALINASSKEEALQIREKIPKYLQFHNLSALA 61  
 DB 111 LEMGFQIDGRSARGCGQOTANDAYAKAINAGSKSQALDVIRELPDYVLFHFNINSNL 170  
 OY 62 AIFDKTPEP 70  
 DB 171 KVFQVPPAP 179

RESULT 11  
 ID VAL1\_ABMV STANDARD; PRT; 355 AA.  
 AC P21947;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 CC AC1.  
 OS Abutilon mosaic virus (isolate West India).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91020984; PubMed=2219703;  
 RA Fischmuth T., Zimmatt G., Jeske H.;  
 RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
 RL as well as eukaryotic features.";  
 RL Virology 178:461-468(1990).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X15983; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; OQCVM1.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR Prodom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 221 228 ATP (POTENTIAL).  
 SQ SEQUENCE 355 AA; 40257 MW; 16A2CA8A63251E95 CRC64;

Query Match 58.5%; Score 210; DB 1; Length 355;  
 Best Local Similarity 57.1%; Pred. No. 1.6e-16;  
 Matches 40; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDGSRAGCGCOTSDAAAEALINASSKEEALQIREKIPKYLQFHNLSAL 60  
 DB 110 TAEWGEFQIDGRSARGCGQOTANDSVAKALINAGVQSALINLEEFQPKDYVLQHNHNSNL 169  
 OY 61 AAFDKTPEP 70  
 DB 170 ERTFAKAP 179

RESULT 12  
 ID VAL1\_TMOV STANDARD; PRT; 361 AA.  
 AC Q06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 CC AL1.  
 OS Tomato mottle virus (isolate Florida) (TMOV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93107858; PubMed=1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 RT isolated from tomatoes in Florida.";  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L14460; AAC32414.1; -;  
 DR PIR: J01870; J01870.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR Prodom: PD000736; Gemini\_AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 222 229 ATP (BY SIMILARITY).  
 SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEBAC6950 CRC64;

Query Match 57.7%; Score 207; DB 1; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 3.5e-16;  
 Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDGSRAGCGCOTSDAAAEALINASSKEEALQIREKIPKYLQFHNLSAL 60  
 DB 110 TLWEDFQIDGRSARGCGQOSANDSVAKALINAGSVALAVIREQPKDYVLQHNHNSNL 169  
 OY 61 AAFDKTPEP 70  
 DB 170 ERTFAKAP 179

RESULT 13  
 ID VAL1\_TYLCV STANDARD; PRT; 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein (C1 protein).  
 CC C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RT with a single genomic component.";  
 RL Virology 185:151-161(1991).

```

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15656; CAA33688.1; -.
DR PIR: D40779; QOCVCL.
DR InterPro: IPR001191; Gemin1_AL1.
DR Pfam: PF00799; Gemin1_AL1.1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemin1_AL1.1.
DR ATP-binding.
KW NP_BIND.
FT SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;
SQ
Query Match 53.8%; Score 193; DB 1; Length 357;
Best Local Similarity 62.3%; Pred. No. 1,4e-14;
Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
OY 4 GGEFVDSRSARGGCGTNDAAAEALNASSKEALQIREKIPKYLEFQPHNLNSALAAI 63
DB 111 FGVSIDERSARGGQGSNDAYAEALNAGSKSEALNLKEKAPDYILQPHNLSSNDRI 170
OY 64 F 64
DB 171 F 171
RESULT 14
VAL1_SLCV STANDARD; PRT; 347 AA.
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE AL1 protein.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OY NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082449; PubMed=1984668;
RA Lazarowitz S.G.; Jazdins I.B.;
RT "Infectivity and complete nucleotide sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype."
RL Virology 180:58-69(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M38183; AAC32410.1; ALT_INIT.
DR PIR: G36785; QOCVSL.
DR InterPro: IPR001191; Gemin1_AL1.
DR Pfam: PF00799; Gemin1_AL1.1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemin1_AL1.1.
DR ATP-binding.
KW NP_BIND.
FT SEQUENCE 347 AA; 39110 MW; AFDABDEDE12210E CRC64;
SQ

```

```

Query Match 29.8%; Score 107; DB 1; Length 347;
Best Local Similarity 34.8%; Pred. No. 7.6e-05;
Matches 23; Conservative 13; Mismatches 26; Indels 4; Gaps 1;
OY 5 GGEFVDSRSARGGCGTNDAAAEALNASSKEALQIREKIPKYLEFQPHNLNSALAAI 64
DB 116 GQYKVSQ-----GSKSNKDDVYHNAAVAGSAGFALDIIRAGPKFTIVYVHNLNVERLF 171
OY 65 DKTPPE 70
DB 172 QKPEP 177
RESULT 15
TBL2_HUMAN STANDARD; PRT; 447 AA.
ID TBL2_HUMAN
AC Q9Y4P3; Q9Q0E2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transducin beta-like 2 protein (WS beta-transducin repeats protein)
DE (WS-betaTRP).
GN TBL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20044626; PubMed=10575226;
RA Perez Jurado L.A.; Wang Y.-K.; Francke U.; Cruces J.;
RT "TBL2, a novel transducin family member in the WBS deletion;
RT characterization of the complete sequence, genomic structure,
RT transcriptional variants and the mouse ortholog."
RL Cytogenet. Cell Genet. 86:277-284(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S.; Well B.; Wellenreuther R.; Gassenhuber J.; Glassl S.;
RA Ansoerge W.; Boecher M.; Bloeker H.; Bauersachs S.; Blum H.;
RA Jauber J.; Duesterhoeft A.; Beyer A.; Koehrer K.; Strack N.;
RA Mewes H.-W.; Ottenwaelder B.; Obermayer B.; Tampe J.; Heubner D.;
RA Wambolt R.; Korn B.; Klein M.; Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs."
RL Genome Res. 11:422-435(2001).
RN [3]
RP SEQUENCE OF 14-447 FROM N.A.
RX MEDLINE=99075645; PubMed=9860302;
RA Meng X.; Lu X.; Li Z.; Green E.D.; Massa H.; Trask B.J.; Morris C.A.;
RA Keating M.T.;
RT "Complete physical map of the common deletion region in Williams
RT syndrome and identification and characterization of three novel
RT genes."
RL Hum. Genet. 103:590-599(1998).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF097484; AAF06823.1; -.
DR EMBL: AF097485; AAF06824.1; -.
DR EMBL: AL080162; CAB45751.1; -.
DR EMBL: AF056183; AAD28083.1; ALT_INIT.
DR MIM: 605842; -.
DR InterPro: IPR001680; WD40.

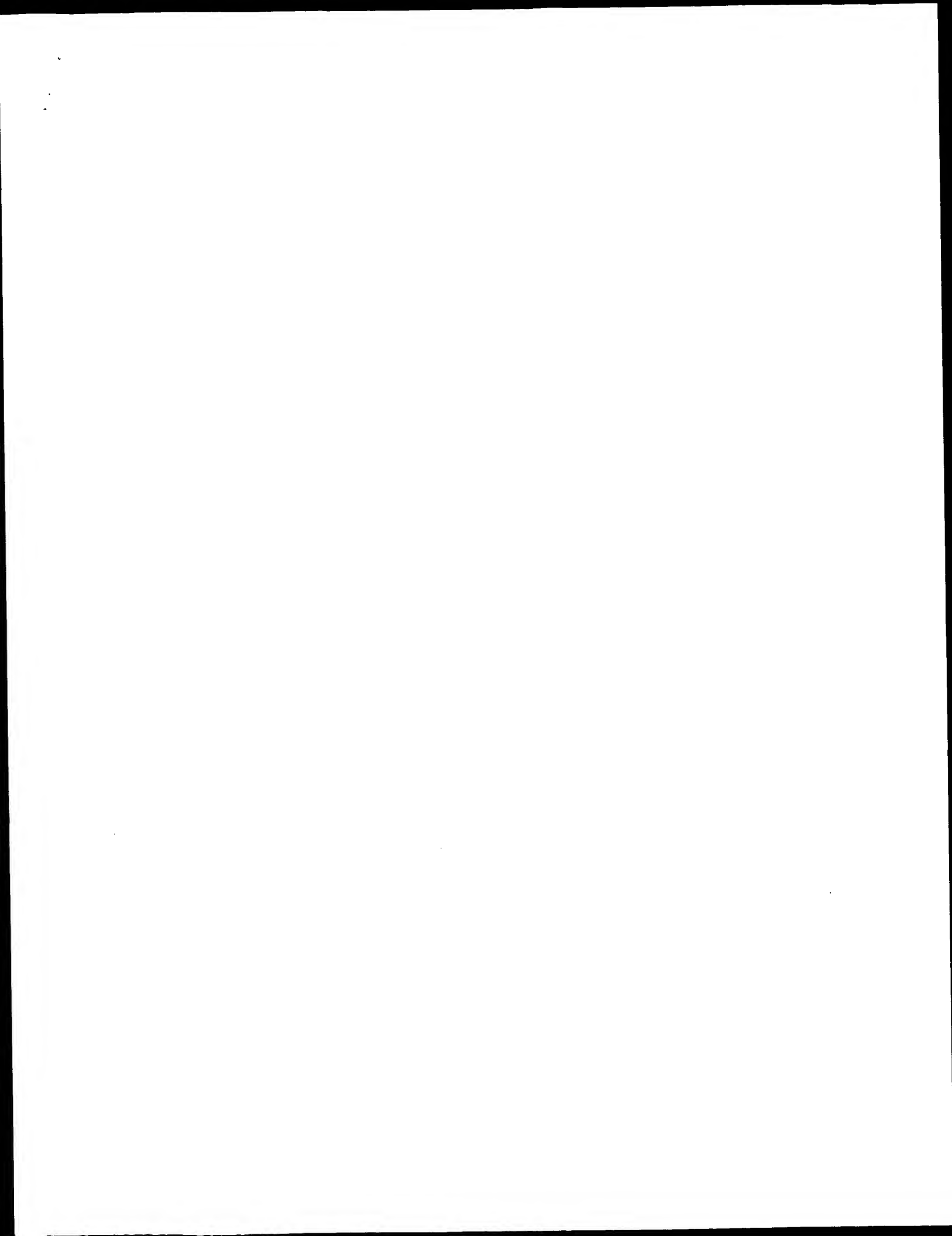
```

DR Pfam: PF00400; WD40; 5  
 DR PRINTS: PRO0320; GPROTEINBRPT.  
 DR SMART: SM00320; WD40; 5.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 3.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat: WD repeat.  
 FT REPEAT 88 127 WD 1.  
 FT REPEAT 134 174 WD 2.  
 FT REPEAT 186 226 WD 3.  
 FT REPEAT 228 267 WD 4.  
 FT REPEAT 277 316 WD 5.  
 FT REPEAT 329 367 WD 6.  
 FT REPEAT 371 409 WD 7.  
 SQ SEQUENCE 447 AA: 49797 MW: B260087E1A71D3F9 CRC64;

Query Match 18.9%; Score 68; DB 1; Length 447;  
 Best Local Similarity 32.7%; Pred. No. 2.7;  
 Matches 16; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

OY 12 RSARGGCGTSDAAAEALNASSKEEALQIREKIPERYLFQFHNLSAL 60  
 Db 38 RSGRPACQKANGFPDPKSSGSKKQKQYQRIKREKPOQHNTTHRLAAL 86

Search completed: October 10, 2002, 02:09:52  
 Job time: 1350 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: October 10, 2002, 01:48:44 ; Search time 69.63 Seconds  
(without alignments)  
96.600 Million cell updates/sec

Title: US-09-289-346A-7

Sequence: 1 TLVWGEFQVDSARSGCOT.....FQFHNLSALAIFFDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 71: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	342	95.3	352	1 OQCVL	ALI protein - toma
2	242	67.4	361	1 OQCVPT	ALI protein - toma
3	229	63.8	358	1 S07594	hypothetical prote
4	222	61.8	362	1 J01887	ALI protein - toma
5	216	60.2	349	1 J02300	replicase - pepper
6	216	60.2	349	2 S31875	ALI protein - pepp
7	215	59.9	359	2 S39211	gene C1 protein -
8	212	59.1	385	2 S28360	ALI protein - beet
9	211	58.8	359	2 S22593	hypothetical prote
10	210	58.5	355	1 OQCVL	ALI protein - toma
11	207	57.7	358	1 J01870	ALI protein - toma
12	205	57.1	351	2 J02327	ALI protein - indi
13	205	57.1	359	2 S39235	gene C1 protein -
14	195	54.3	360	2 S59885	replication-associ
15	193	53.8	357	1 OQCVL	ALI protein - toma
16	193	53.8	357	1 OQCVL	ALI protein (clone
17	120	33.4	131	2 S45059	ALI protein - squa
18	107	29.8	347	1 OQCVS1	Fc gamma (19c) rec
19	70	19.5	587	2 J01419	hypothetical prote
20	68	18.9	447	2 T12544	leucine aminopepti
21	64.5	18.0	491	2 AE1836	conserved hypotet
22	64	17.8	316	2 C82085	probably O-sialogl
23	63.5	17.7	340	2 B97011	kinesin-like prote
24	63.5	17.7	1229	2 T48959	voltage-dependent
25	63.5	17.7	1610	2 A46227	voltage-dependent
26	63.5	17.7	2164	2 JH0422	calcium channel al
27	63.5	17.7	2181	2 A38198	calcium channel al
28	63.5	17.7	2203	2 T42742	voltage-dependent
29	62	17.3	154	2 AD3475	nitrogen regulator

30	62	17.3	245	2 S76632	hypothetical prote
31	62	17.3	714	2 C95382	probable ferric h
32	61	17.0	1502	1 RGHVH1	Cycl/CYP3 transcr
33	60.5	16.9	276	2 AC2341	hypothetical prote
34	60.5	16.9	840	2 T36175	probable large ATP
35	59.5	16.6	208	2 E97702	hypothetical prote
36	59.5	16.6	224	2 A10474	pro-sigma-E proces
37	59.5	16.6	408	2 A99238	conserved hypotet
38	59.5	16.6	446	1 S17653	dyein beta heavy
39	59	16.4	338	2 AG3582	iron(III)-binding
40	59	16.4	345	2 AD3024	hypothetical prote
41	59	16.4	359	2 E98260	periplasmic bindin
42	59	16.4	370	2 T34050	hypothetical prote
43	59	16.4	384	1 W2ML41	E2 protein - human
44	59	16.4	2459	1 AF2136	peptide synthetase
45	58.5	16.3	136	2 T22240	hypothetical prote

## ALIGNMENTS

## RESULT 1

OQCVL 1  
ALI protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 95.3%, Score 342, DB 1, Length 352;

Best Local Similarity 95.7%; Pred. No. 2.6e-31;

Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSARSGCOTSDNAABALNASSKEALQITREKIPKYLFOFHNLSAL 60

DB 111 TLVWGEFQVDSARSGCOTSDNAABALNASSKEALQITREKIPKYLFOFHNLSNL 170

OY 61 AAIFDKTPPEP 70

DB 171 DRIFDKTPPEP 180

RESULT 2

OQCVPT 2

ALI protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C:Accession: J00364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato y

A:Reference number: J00362; MUID:91311403

A:Accession: J00364

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-361 <COU>

A:Cross-references: GB:D00940; NID:9222458; PIDN:BAA00782.1; PID:9222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 67.4%; Score 242; DB 1; Length 361;  
 Best Local Similarity 66.7%; Pred. No. 7e-20;  
 Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLSAL 60  
 Db 110 TLEWGEFQIDGRSARGGQGSANDYAKALNSGSKSFALNVLRELAKPKYVLOFHNLSNL 169  
 OY 61 AAIKDKTPEP 69  
 Db 170 DRIFPPLEVVSPFLSSFDVRPE 194

RESULT 3  
 S07594  
 hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)  
 C:Species: cassava latent virus  
 C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
 C:Accession: S07594  
 R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.  
 Nucleic Acids Res. 18, 197-198, 1990  
 A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava  
 A:Reference number: S07590; MUID:90174930  
 A:Accession: S07594  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <MOR>  
 A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376  
 C:Genetics:  
 A:Map position: segment DNA1  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 63.8%; Score 229; DB 2; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 2.1e-18;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLSAL 60  
 Db 109 TVEWGEFQIDGRSARGGQGSANDYAKALNSGSKSFALNVLRELAKPKYVLOFHNLSNL 168  
 OY 61 AAIKDKTPEP 70  
 Db 169 DRIFQTPPEP 178

RESULT 4  
 JQ1887  
 AL1 protein - tomato yellow leaf curl virus (strain Australia)  
 N:Alternate names: CL protein  
 C:Species: tomato yellow leaf curl virus  
 C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: JQ1887  
 R:Dry, I.B.; Ridgen, J.E.; Krahe, L.R.; Mullineaux, P.M.; Rezaian, M.A.  
 J. Gen. Virol. 74, 147-151, 1993  
 A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.  
 A:Reference number: JQ1885; MUID:93139778  
 A:Accession: JQ1887  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-362 <DRY>  
 A:Cross-references: GB:S53251  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.8%; Score 222; DB 1; Length 362;  
 Best Local Similarity 52.9%; Pred. No. 1.4e-17;  
 Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

OY 1 TLVWGEFQVDSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLSAL 60  
 Db 110 TLEWGEFQIDGRSARGGQGSANDYAKALNSGSKSFALNVLRELAKPKYVLOFHNLSNL 169

Db 110 TLEWGEFQIDGRSARGGQGSANDYAKALNTGSKSFALNVLRELAKPKYVLOFHNLSNL 169  
 OY 61 AAI-----FDKTPPE 69  
 Db 170 DRIFPPLEVVSPFLSSFDVRPE 194

RESULT 5  
 JQ2300  
 replicase - pepper huasteco virus (component A)  
 N:Alternate names: ORF AL1 protein  
 C:Species: pepper huasteco virus  
 C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999  
 C:Accession: JQ2300  
 R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante, J. Gen. Virol. 74, 2225-2231, 1993  
 A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b  
 A:Reference number: JQ2299; MUID:94015007  
 A:Accession: JQ2300  
 A:Molecule type: DNA  
 A:Residues: 1-349 <TOR>  
 A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.2%; Score 216; DB 2; Length 349;  
 Best Local Similarity 58.6%; Pred. No. 6.3e-17;  
 Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLSAL 60  
 Db 110 TVEWGEFQIDGRSARGGQGSANDYAKALNSAEEALQITREKIPKYLQFPHNLSNA 169  
 OY 61 AAIKDKTPEP 70  
 Db 170 NRIFQTPPEP 179

RESULT 6  
 S31875  
 AL1 protein - pepper rizado amarillo virus  
 C:Species: pepper rizado amarillo virus  
 C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999  
 C:Accession: S31875  
 R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera  
 submitted to the EMBL data library, February 1993  
 A:Description: Complete nucleotide sequence of pepper huasteco virus: analysis and co  
 A:Reference number: S31872  
 A:Accession: S31875  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-349 <TOR>  
 A:Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025  
 C:Note: the source is designated as pepper huasteco virus  
 C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.2%; Score 216; DB 2; Length 349;  
 Best Local Similarity 58.6%; Pred. No. 6.3e-17;  
 Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLSAL 60  
 Db 110 TVEWGEFQIDGRSARGGQGSANDYAKALNSAEEALQITREKIPKYLQFPHNLSNA 169  
 OY 61 AAIKDKTPEP 70  
 Db 170 NRIFQTPPEP 179

RESULT 7  
 S39211  
 gene CL protein - tomato yellow leaf curl virus

Query Match	57.7%	Score 207;	DB 1;	Length 358;
Best Local Similarity	54.3%;	Pred. No. 6.9e-16;		
Matches	38;	Conservative	15;	Mismatches 17;
				Indels 0;
				Gaps 0;

QY 1 TLWGEFQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 60  
 DB 107 TLMGFEQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 166  
 QY 61 AATFDKTPPEP 70  
 DB 167 ERIFAKAPPEP 176

## RESULT 12

JQ2327  
 A: protein - Indian cassava mosaic virus  
 N: Alternate names: replication-associated protein  
 C: Species: Indian cassava mosaic virus  
 C: Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
 C: Accession: JQ2327; S35883  
 R: Hong, Y.G.; Robinson, D.J.; Harrison, B.D.  
 J. Gen. Virol. 74, 2437-2443, 1993  
 A: Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted geminiviruses  
 A: Reference number: JQ2326; MUID:94065670  
 A: Accession: JQ2327  
 A: Molecule type: DNA  
 A: Residues: 1-351 <HON>  
 A: Cross-references: EMBL:Z24758; NID:q395351; PIDN:CAA80891.1; PID:g584046  
 C: Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.1%; Score 205; DB 2; Length 351;  
 Best Local Similarity 59.7%; Pred. No. 1.1e-15;  
 Matches 40; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 4 WGEFQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 63  
 DB 113 TLMGFEQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 172  
 QY 64 FDKTPPEP 70  
 DB 173 FTKEPPP 179

## RESULT 13

S39235  
 gene C1 protein - tomato yellow leaf curl virus  
 C: Species: tomato yellow leaf curl virus  
 C: Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
 C: Accession: S39235  
 R: Crespi, S.; Norris, E.; Vaira, A.; Bosco, D.; Accotto, G.  
 submitted to the EMBL data library, December 1993  
 A: Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.  
 A: Reference number: S39233  
 A: Accession: S39235  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-359 <CRE>  
 A: Cross-references: EMBL:Z28390; NID:q1041671; PID:g1334964  
 C: Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.1%; Score 205; DB 2; Length 359;  
 Best Local Similarity 55.1%; Pred. No. 1.2e-15;  
 Matches 38; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 2 TLMGFEQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 61  
 DB 111 TLMGFEQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 170  
 QY 62 AATFDKTPPEP 70  
 DB 171 KVFQVPPAP 179

## RESULT 14

S59885  
 replication-associated protein C1 - tomato yellow leaf curl virus  
 C: Species: tomato yellow leaf curl virus  
 C: Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C: Accession: S59885  
 R: Hong, Y.; Harrison, B.D.  
 submitted to the EMBL data library, February 1995  
 A: Description: Nucleotide sequences from tomato leaf curl viruses from different countries  
 A: Reference number: S58346  
 A: Accession: S59885  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-360 <HON>  
 A: Cross-references: EMBL:Z48182; NID:q944838; PIDN:CAA8229.1; PID:g974211  
 C: Superfamily: tomato golden mosaic virus AL1 protein

Query Match 54.3%; Score 195; DB 2; Length 360;  
 Best Local Similarity 57.6%; Pred. No. 1.6e-14;  
 Matches 38; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 4 WGEFQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 63  
 DB 113 TLMGFEQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 172  
 QY 64 FDKTPPEP 69  
 DB 173 FTKEPPP 178

## RESULT 15

QOCVC1  
 A1 protein - tomato yellow leaf curl virus  
 N: Alternate names: C1 protein  
 C: Species: tomato yellow leaf curl virus  
 C: Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C: Accession: D40779  
 R: Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A: Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 A: Reference number: M40779; MUID:92024070  
 A: Accession: D40779  
 A: Status: translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-357 <NAV>  
 A: Cross-references: GB:X15656; NID:q62204; PIDN:CAA33688.1; PID:g62207  
 C: Superfamily: tomato golden mosaic virus AL1 protein

Query Match 53.8%; Score 193; DB 1; Length 357;  
 Best Local Similarity 62.3%; Pred. No. 2.7e-14;  
 Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGEFQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 63  
 DB 111 TLMGFEQVDSARAGCGCOTSNDAAEALNASSKEEALQIREKIPKYLFOFHNLNSALA 170  
 QY 64 F 64  
 DB 171 F 171

Search completed: October 10, 2002, 01:48:45  
 Job time: 563 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds  
(without alignments)  
58.394 Million cell updates/sec

Title: US-09-289-346A-7

Sequence: 1 TLWGEFQVDGSAAGCGQT.....FGPHNLNALAIIDKTPPEP 70

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: A\_Geneseq\_032802.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	70	21	AA18683
2	342	95.3	70	21	AA18677
3	342	95.3	356	21	AA18687
4	334	93.0	70	21	AA18685
5	331	92.2	70	21	AA18688
6	330	91.9	70	21	AA18692
7	328	91.4	70	21	AA18684
8	328	91.4	70	21	AA18690
9	327	91.1	70	21	AA18678
10	327	91.1	70	21	AA18686
11	326	90.8	70	21	AA18689

12	324	90.3	70	21	AA18680	Mutant peptide der
13	324	90.3	70	21	AA18691	Mutant peptide der
14	321	89.4	70	21	AA18681	Mutant peptide der
15	318	88.6	70	21	AA18682	Mutant peptide der
16	310	86.4	70	21	AA18679	Mutant peptide der
17	213	59.3	353	18	AAW34338	Bean golden mosaic
18	213	59.3	353	18	AAW34332	Bean golden mosaic
19	213	59.3	353	18	AAW34333	Bean golden mosaic
20	213	59.3	353	18	AAW34334	Bean golden mosaic
21	213	59.3	353	18	AAW34335	Bean golden mosaic
22	211	58.8	353	8	AAW70407	ORF 4 gene product
23	211	58.8	359	17	AAW88870	Sardinian tomato y
24	211	58.8	359	17	AAW88871	Sardinian tomato y
25	211	58.8	359	17	AAW88872	Sardinian tomato y
26	211	58.8	359	18	AAW34336	Tomato mottle viru
27	211	58.8	361	18	AAW34324	Tomato mottle viru
28	211	58.8	361	18	AAW34325	Tomato mottle viru
29	211	58.8	361	18	AAW34326	Tomato mottle viru
30	208.5	58.1	361	8	AAW70562	Product of ORF 4 f
31	194	54.0	362	19	AAW56495	Tobacco leaf curl
32	193	53.8	357	18	AAW34329	Tomato yellow leaf
33	193	53.8	357	18	AAW34330	Tomato yellow leaf
34	193	53.8	357	18	AAW34331	Tomato yellow leaf
35	185	51.5	357	18	AAW34337	Tomato yellow leaf
36	68	18.9	447	21	AAW99659	Human GTPase assoc
37	68	18.9	447	21	AAW87089	Human secreted pro
38	68	18.9	447	21	AAW66744	Membrane-bound pro
39	68	18.9	447	21	AAW50947	Human adult aorta
40	68	18.9	447	22	AAW29150	Human PRO polypept
41	68	18.9	447	22	AAW93267	Human polypeptide,
42	68	18.9	447	22	AAW50666	Human gene 26 enco
43	68	18.9	447	22	AAW65267	Human PRO1125 (UNQ
44	68	18.9	456	21	AAW87190	Human secreted pro
45	68	18.9	456	22	AAW06167	Human gene 26 enco

#### ALIGNMENTS

RESULT 1	AA18683	standard; peptide; 70 AA.
XX	AA18683	
AC	AA18683	
XX	22-JAN-2001	(first entry)
DT		
XX		
DE		Mutant peptide derived from amino acids 110-179 of Rep (All) protein.
KW		Geminivirus; replication protein; Rep protein; All; transgenic plant;
KM		ribosome binding region; resistance; geminivirus infection.
XX		
OS		Synthetic.
OS		Tomato golden mosaic virus.
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 59	/note= "wild type residue replaced with Ala"
FT	Misc-difference 61	/note= "wild type residue replaced with Ala"
FT	Misc-difference 62	/note= "wild type residue replaced with Ala"
FT		
XX	MO200054573-Al.	
XX		
XX	21-SEP-2000.	
PD		
XX		
PF	15-MAR-2000.	2000MO-US06759.
XX		
XX		
PR	18-MAR-1999.	99US-0125004.
PR	09-APR-1999.	99US-0289346.
XX		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI: 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PS  
 XX Claim 53; Page 45; 73pp; English.  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX Sequence 70 AA;  
 SQ  
 Query Match 100.0%; Score 359; DB 21; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-39;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TLVWGEFVQDGRSARGCGQTNDAAAFALNASSKEEALQIIREKIPKYLFOFHNLSAL 60  
 Db 1 TLVWGEFVQDGRSARGCGQTNDAAAFALNASSKEEALQIIREKIPKYLFOFHNLSAL 60  
 QY 61 AAIIDKTPPEP 70  
 Db 61 aatfdktpep 70

RESULT 2  
 AAB18677 standard; peptide: 70 AA.  
 AC AAB18677;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI: 2000-618851/59.  
 XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PS  
 XX Disclosure; Page 18; 73pp; English.  
 XX The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 CC  
 XX Sequence 70 AA;  
 SQ  
 Query Match 95.3%; Score 342; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 1.2e-36;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 TLVWGEFVQDGRSARGCGQTNDAAAFALNASSKEEALQIIREKIPKYLFOFHNLSAL 60  
 Db 1 TLVWGEFVQDGRSARGCGQTNDAAAFALNASSKEEALQIIREKIPKYLFOFHNLSAL 60  
 QY 61 AAIIDKTPPEP 70  
 Db 61 dritfdktpep 70

RESULT 3  
 AAB18687 standard; peptide: 356 AA.  
 AC AAB18687;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Amino acid sequence of a geminivirus replication protein of TGMV.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 354 /note= "unspecified amino acid"  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI: 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
XX  
PS Disclosure; Page 47-48; 73pp; English.  
XX  
CC The present sequence represents a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
XX  
SQ Sequence 356 AA:  
  
Query Match 95.3%; Score 342; DB 21; Length 356;  
Best Local Similarity 95.7%; Pred. No. 9.3e-36;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 TLVGEFVDGSRAGCCGCTSDAAAEALNMSKKEALQIREKIPKYLFOFHNLSAL 60  
DB 110 TLVGEFVGYGSRAGCGTSDAAAEALNMSKKEALQIREKIPKYLFGFHNLSNL 169  
QY 61 AATFDKTPPEP 70  
DB 170 drlfdktpep 179  
  
RESULT 4  
AAB18685  
ID AAB18685 standard; peptide; 70 AA.  
XX  
AC AAB18685;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
OS Tomato golden mosaic virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 10 /note= "wild type residue replaced with Ala"  
FT  
XX  
XX WO200054573-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000MO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX WPI; 2000-618851/59.  
XX  
PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
XX  
PS Claim 53; Page 46; 73pp; English.  
XX  
CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA:  
  
Query Match 93.0%; Score 334; DB 21; Length 70;  
Best Local Similarity 94.3%; Pred. No. 1.3e-35;  
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 TLVGEFVDGSRAGCCGCTSDAAAEALNMSKKEALQIREKIPKYLFOFHNLSAL 60  
DB 1 TLVGEFVGYGSRAGCGTSDAAAEALNMSKKEALQIREKIPKYLFGFHNLSNL 169  
QY 61 AATFDKTPPEP 70  
DB 61 drlfdktpep 70  
  
RESULT 5  
AAB18688  
ID AAB18688 standard; peptide; 70 AA.  
XX  
AC AAB18688;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
OS Tomato golden mosaic virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 19 /note= "wild type residue replaced with Ala"  
FT  
XX  
XX WO200054573-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000MO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX WPI; 2000-618851/59.  
XX  
PT Transgenic plants with increased resistance to geminivirus infection

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PS  
 XX Disclosure: Page 48; 73pp; English.

XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.

XX Sequence 70 AA:

Query Match 92.2%; Score 331; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 3, 1e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGSRAGCGCQTSDNDAAEALNASSKEALQIIREKIPKYLQFPHNLNSAL 60  
 Db 1 TLVWGEFVDGSRAGCGCQTSDNDAAEALNASSKEALQIIREKIPKYLQFPHNLNSAL 60  
 OY 61 AAFDKRPEP 70  
 Db 61 drlfdktpep 70

Db 61 drlfdktpep 70

RESULT 6  
 AAB18692  
 ID AAB18692 standard; peptide: 70 AA.

XX AAB18692:

XX 22-JAN-2001 (first entry)

XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

XX ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

XX Tomato golden mosaic virus.

XX Key Location/Qualifiers

XX MISC-difference 66 /note= "wild type residue replaced with Ala"

XX MISC-difference 69 /note= "wild type residue replaced with Ala"

XX WO200054573-A1.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.

XX 09-APR-1999; 99US-0289346.

XX (UYN-C-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PS  
 XX Disclosure: Page 50; 73pp; English.

XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.

XX Sequence 70 AA:

Query Match 91.9%; Score 330; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 4, 1e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGSRAGCGCQTSDNDAAEALNASSKEALQIIREKIPKYLQFPHNLNSAL 60  
 Db 1 TLVWGEFVDGSRAGCGCQTSDNDAAEALNASSKEALQIIREKIPKYLQFPHNLNSAL 60  
 OY 61 AAFDKRPEP 70  
 Db 61 drlfdktpep 70

Db 61 drlfdktpep 70

RESULT 7  
 AAB18684  
 ID AAB18684 standard; peptide: 70 AA.

XX AAB18684:

XX 22-JAN-2001 (first entry)

XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

XX ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

XX Tomato golden mosaic virus.

XX Key Location/Qualifiers

XX MISC-difference 7 /note= "wild type residue replaced with Ala"

XX MISC-difference 8 /note= "wild type residue replaced with Ala"

XX WO200054573-A1.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.

XX 09-APR-1999; 99US-0289346.

XX (UYN-C-) UNIV NORTH CAROLINA STATE.



XX	PI	Hanley-Bowdoin L, Orozco BM, Kong L.
XX	DR	WP1; 2000-618851/59.
XX	PT	Transgenic plants with increased resistance to geminivirus infection
XX	PT	comprise a nucleic acid construct containing a nucleic acid sequence
XX	PT	encoding a mutant AL1 protein with a mutation in the Rb binding region
XX	PS	-
XX	XX	Claim 52; Page 45; 73pp; English.
CC	CC	The present sequence represents a mutant peptide, derived from a
CC	CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC	CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC	CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC	CC	protein are used to produce transgenic plants. The mutation in AL1 is
CC	CC	present in a ribosome binding region, and expression of mutant AL1
CC	CC	protein imparts increased resistance to geminivirus infection in the
CC	CC	plant. Mutant AL1 proteins are useful for producing plants having
CC	CC	increased resistance or reduced sensitivity to a geminivirus such as
CC	CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC	CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC	CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC	CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC	CC	virus, cotton leaf curl virus or beet curly top virus.
XX	XX	
XX	Sequence	70 AA;

```

PF 15-MAR-2000; 2000MO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
XX Claim 53; Page 42-43; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyzes cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 91.1%; Score 327; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1e-34;
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDSRGSGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFPFHNLSAL 60
DB 1 tlwgefgvdsrgsgcqtndalllealnasskeaalqiiirekipekyltqfhnlnsl 60
QY 61 AAIPEKTPPEP 70
DB 61 drifdktppep 70
Db
RESULT 10
AAB18686
ID AAB18686 standard; peptide; 70 AA.
XX
XX AAB18686;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 24 /note= "wild type residue replaced with Leu"
XX
XX Misc-difference 25 /note= "wild type residue replaced with Leu"
XX
XX Misc-difference 26 /note= "wild type residue replaced with Leu"
XX
XX
XX

```

```

PN WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000MO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
XX Claim 53; Page 46; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyzes cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 91.1%; Score 327; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1e-34;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDSRGSGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFPFHNLSAL 60
DB 1 tlwgefgvdsrgsgcqtndalllealnasskeaalqiiirekipekyltqfhnlnsl 60
QY 61 AAIPEKTPPEP 70
DB 61 drifdktppep 70
Db
RESULT 11
AAB18689
ID AAB18689 standard; peptide; 70 AA.
XX
XX AAB18689;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 22 /note= "wild type residue replaced with Ala"
XX
XX
XX

```

FT /note= "wild type residue replaced with Ala"  
 XX MO200054573-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI  
 XX WPI; 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT -  
 PS Disclosure; Page 48-49; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX Sequence 70 AA;  
 SQ

Query Match 90.8%; Score 326; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 1.4e-34;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGCGCOTSDNAAEALNASSKEEALQIREKIPKYLPOFHNINSL 60  
 DB |||||||||||||||||| |||||||||||||||||| ||||||||||||||||  
 1 tlwgefgydgsargcgctsaataaaalnasskeeaqlirekipekylfghninsnl 60

QY 61 AAIIDKTPPP 70  
 DB ||||||||  
 61 drlfdktpep 70

RESULT 12  
 AAB18680  
 ID AAB18680 standard; peptide: 70 AA.  
 AC AAB18680;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 XX Tomato golden mosaic virus.  
 XX  
 XX Key Location/Qualifiers  
 FH MISC-difference 42  
 FT

FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 43  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 44  
 FT /note= "wild type residue replaced with Ala"  
 PN  
 XX MO200054573-A1.  
 PD  
 XX 21-SEP-2000.  
 PF  
 XX 15-MAR-2000; 2000WO-US06759.  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI  
 XX WPI; 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT -  
 PS Claim 52; Page 43-44; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX Sequence 70 AA;  
 SQ

Query Match 90.3%; Score 324; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 2.5e-34;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGCGCOTSDNAAEALNASSKEEALQIREKIPKYLPOFHNINSL 60  
 DB |||||||||||||||||| |||||||||||||||||| ||||||||||||||||  
 1 tlwgefgydgsargcgctsaataaaalnasskeeaqliaaapekylfghninsnl 60

QY 61 AAIIDKTPPP 70  
 DB ||||||||  
 61 drlfdktpep 70

RESULT 13  
 AAB18691  
 ID AAB18691 standard; peptide: 70 AA.  
 AC AAB18691;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 XX

OS		Tomato golden mosaic virus.
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 34	/note= "Wild type residue replaced with Ala"
FT	Misc-difference 35	/note= "Wild type residue replaced with Ala"
FT	Misc-difference 36	/note= "Wild type residue replaced with Ala"
FT		
FN	WO200054573-A1.	
XX		
PD	21-SEP-2000.	
XX		
PF	15-MAR-2000; 2000WO-US06759.	
XX		
PR	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
XX		
PI	Hanley-Bowdoin L, Orozco BM, Kong L;	
DR	WPI; 2000-618851/59.	
XX		
PT	Transgenic plants with increased resistance to geminivirus infection	
PT	comprise a nucleic acid construct containing a nucleic acid sequence	
PT	encoding a mutant AL1 protein with a mutation in the Rb binding region	
PT	-	
XX		
PS	Disclosure; Page 49; 73pp; English.	
XX		
CC	The present sequence represents a mutant peptide, derived from a	
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds	
CC	double-stranded DNA, catalyses cleavage and ligation of single-strands	
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1	
CC	protein are used to produce transgenic plants. The mutation in AL1 is	
CC	present in a ribosome binding region, and expression of mutant AL1	
CC	protein imparts increased resistance to geminivirus infection in the	
CC	plant. Mutant AL1 proteins are useful for producing plants having	
CC	increased resistance or reduced sensitivity to a geminivirus such as	
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl	
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian	
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic	
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper	
CC	virus, cotton leaf curl virus or beet curly top virus..	
XX		
SO	Sequence 70 AA:	
	Query Match 90.3%; Score 324; DB 21; Length 70;	
	Best Local Similarity 91.4%; Pred. No. 2.5e-34;	
	Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
OY	1 TLVWEFOVDGSAAGCGCOTSDNDAAEALNASKSEALQITREKIPEKTLPOFHNINSATL 60	
Db	1 tlwgefgydgysargcgctsndaaaaaaalqlirekipekyllqfmlnlnsl 60	
OY	61 AAIIDTRPDP 70	
Db	61 drifdkrpep 70	
	RESULT 14	
ID	AAB18681	
XX	AAB18681 standard; peptide: 70 AA.	
AC	AAB18681;	
XX		
DT	22-JAN-2001 (first entry)	
XX		
DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.	

KM GeminiVirus; replication protein; Rep protein; ALI; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
OS Synthetic.  
XX Tomato golden mosaic virus.

FH Key Location/Qualifiers  
FT Misc-difference 47 /note= "wild type residue replaced with Ala"  
FT Misc-difference 48 /note= "wild type residue replaced with Ala"  
FT Misc-difference 49 /note= "wild type residue replaced with Ala"  
FT WO200054573-AL.  
PN 21-SEP-2000.  
PD 15-MAR-2000; 2000WO-US06759.  
PF 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
PA Hanley-Bowdoin L, Orozco BM, Kong L;  
PI WPT: 2000-618851/59.  
DR Transgenic plants with increased resistance to geminivirus infection  
XX comprise a nucleic acid construct containing a nucleic acid sequence  
XX encoding a mutant ALI protein with a mutation in the Rb binding region

Pt -  
PS Claim 52; Page 44; 73pp; English.  
XX

The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as ALI. ALI binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the ALI  
CC protein are used to produce transgenic plants. The mutation in ALI is  
CC present in a ribosome binding region, and expression of mutant ALI  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant ALI proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX

SQ Sequence 70 AA:  
Query Match 89.4%; Score 321; DB 21; Length 70;  
Best Local Similarity 91.4%; Pred. No. 6e-34;  
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 TLWGEFVDGSGAGCGCOTSDNDAAEALINASSKKEAQLIIREIKPEKYLFQFHMINLSAL 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |  
Db 1 tlwgeifvdgsaragcgctsndaaaealinasskkeaalqlitrekipeaalfghnlmsl 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |  
OY 61 AAIPEKTPEP 70  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |  
Db 61 drlfktpep 70

RESULT 15  
AAAB18682 standard; peptide: 70 AA.  
ID ABAB18682 standard; peptide: 70 AA.  
XX  
AC ABAB18682;  
XX

Job time: 528 sec

DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Msc-difference 52  
 FT /note= "wild type residue replaced with Ala"  
 FT Msc-difference 54  
 FT /note= "wild type residue replaced with Ala"  
 FT Msc-difference 55  
 FT /note= "wild type residue replaced with Ala"  
 XX  
 XX WO200054573-A1.  
 XX  
 XX 21-SEP-2000.  
 PD  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI  
 XX WPI; 2000-618851/59.  
 DR  
 XX  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 53: Page 44-45; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX  
 SQ Sequence 70 AA;  
 Query Match 88.6%; Score 318; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 1.5e-33;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVGRSARSGCOTSDNAAFALNASSKEALQIRREKIPKYLFOFHNLNSAL 60  
 DB 1 TLVWGEFQVGRSARSGCOTSDNAAFALNASSKEALQIRREKIPKYLfafaalnsnl 60  
 QY 61 AALFDKTPPEP 70  
 DB 61 dritdktppep 70



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:02:00 ; Search time 118.38 Seconds  
(without alignments)  
102.295 Million cell updates/sec

Title: US-09-289-346a-8

Perfect score: 361  
Sequence: 1 TLWGEAAVDSGRSARCCGT.....TGFHNLNSLDRIKTPPEP 70

Scoring table: BLAST62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL19:\*\*

- 1: sp-archaea:\*\*
- 2: sp-bacteria:\*\*
- 3: sp-fungi:\*\*
- 4: sp-human:\*\*
- 5: sp-invertebrate:\*\*
- 6: sp-mammal:\*\*
- 7: sp-mhc:\*\*
- 8: sp-organelle:\*\*
- 9: sp-phage:\*\*
- 10: sp-plant:\*\*
- 11: sp-rodent:\*\*
- 12: sp-virus:\*\*
- 13: sp-vertebrate:\*\*
- 14: sp-unclassified:\*\*
- 15: sp-rv:\*\*
- 16: sp-bacteriophage:\*\*
- 17: sp-archaeal:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	301	83.4	351	12	Q91R10 tomato seve
2	299	82.8	352	12	Q9E000 tomato rugo
3	282	78.1	226	12	Q09727 leonurus mo
4	282	78.1	226	12	Q9WHF6 tomato mild
5	280	77.6	361	12	Q67574 bean golden
6	278	77.0	225	12	Q9QD81 cowpea gold
7	278	77.0	314	12	Q9ELR8 sweet potat
8	278	77.0	364	12	Q9G555 sweet potat
9	275	76.2	185	12	Q98693 sida golden
10	267	74.0	149	12	Q98975 macroptilin
11	267	74.0	233	12	Q9YL44 macroptilin
12	261	72.3	190	12	Q39180 geminivirid
13	258	71.5	190	12	Q92089 tobacco lea
14	258	71.5	190	12	Q92084 tobacco lea
15	257	71.2	190	12	Q9W827 tobacco lea
16	257	71.2	208	12	Q920C4 tobacco lea

17	257	71.2	359	12	Q91M88	Q91M88 tobacco lea
18	255	70.6	208	12	Q920C0	Q920C0 tobacco lea
19	255	70.6	208	12	Q920B8	Q920B8 tobacco lea
20	253	70.1	208	12	Q920B6	Q920B6 tobacco lea
21	252	69.8	203	12	Q92083	Q92083 tobacco lea
22	252	69.8	363	12	Q92083	Q92083 tobacco lea
23	247	68.4	359	12	Q91B86	Q91B86 ageratum ye
24	247	68.4	360	12	Q9DX10	Q9DX10 ageratum ye
25	245	67.9	208	12	Q920A0	Q920A0 tobacco lea
26	245	67.9	363	12	Q92705	Q92705 cotton leaf
27	245	67.9	363	12	Q92719	Q92719 cotton leaf
28	244	67.6	190	12	Q92086	Q92086 tobacco lea
29	244	67.6	208	12	Q920C6	Q920C6 tobacco lea
30	243	67.3	349	12	Q98888	Q98888 tomato pseu
31	243	67.3	362	12	Q56816	Q56816 chayote mos
32	243	67.3	359	12	Q912V4	Q912V4 tomato yell
33	242	67.0	359	12	Q912V2	Q912V2 tomato yell
34	242	67.0	359	12	Q912V7	Q912V7 tomato yell
35	242	67.0	359	12	Q912V7	Q912V7 tomato yell
36	241	66.8	348	12	Q911W5	Q911W5 macroptilin
37	241	66.8	353	12	Q72692	Q72692 beet curly
38	241	66.8	354	12	Q65438	Q65438 beet curly
39	241	66.8	359	12	Q88942	Q88942 tomato yell
40	240	66.5	190	12	Q920A7	Q920A7 tobacco lea
41	240	66.5	362	12	Q91N48	Q91N48 okra enatio
42	239	66.2	307	12	Q91ET7	Q91ET7 cotton leaf
43	239	66.2	361	12	Q72723	Q72723 cotton leaf
44	238	65.9	231	12	Q96620	Q96620 african tom
45	238	65.9	354	12	Q91N42	Q91N42 south afric

## ALIGNMENTS

RESULT 1  
ID Q91R10 PRELIMINARY: PRT: 351 AA.  
AC Q91R10;  
DT 01-DEC-2001 (TREMREL. 19, Created)  
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)  
DE 01-DEC-2001 (TREMREL. 19, Last annotation update)  
DE REP PROTEIN.  
GN AC1.  
OS Tomato severe rugose virus.  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=158463;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MINIS GERALS;  
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;  
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting  
begomovirus, Tomato severe rugose virus, in Brazil."  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF029750; AAK50357.1; -; 87F937A4F873B6CF CXC64;  
SQ SEQUENCE 351 AA: 40122 MW: 87F937A4F873B6CF CXC64;

Query Match 83.4%; Score 301; DB 12; Length 351;  
Best Local Similarity 80.0%; Pred. No. 5.6e-26;  
Matches 56; Conservatively 7; Mismatches 7; Indels 0; Gaps 0;  
QY 1 TLWGEAAVDSGRSARCCGTSDAAEAALNASKPEFALDIIEKIPKYLTFGFHNLNSNL 60  
Db 111 TLWGEAAVDSGRSARCCGTANDAAEAALNASKPEFALDIIEKIPKYLTFGFHNLNSNL 170  
QY 61 DRIFRATPEP 70  
Db 171 DRIFRATPEP 180  
RESULT 2  
ID Q9E000 PRELIMINARY: PRT: 352 AA.

AC Q9E000;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-JUN-2001 (TREMBLrel. 16, last sequence update)  
 DE REPLICATION-ASSOCIATED PROTEIN.  
 GN ACL  
 OS Tomato rugose mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=134599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,  
 RA Zamolli E.M., Zerbini F.M.;  
 RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
 (TRMV), a Begomovirus isolated from Tomatoes at Triangulo Mineiro,  
 RT Minas Gerais, Brazil."  
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF291705; AAC15546.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 SQ SEQUENCE 352 AA; 40012 MW; 47CD5838E24D613 CRC64;

Query Match 82.8%; Score 299; DB 12; Length 352;  
 Best Local Similarity 80.0%; Pred. No. 9.5e-26;  
 Matches 56; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVWGEAAVDSRGSGCOTSDAAAEALNASSKEEALQIREKIPKEYLFOFHNLNSL 60  
 DB 111 TLVWGEFQIDGSRGSGCOTANDAAAEALNASSKEEALQIREKIPKEYLFOFHNLNSL 170  
 QY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 3  
 009727  
 ID 009727 PRELIMINARY; PRT; 226 AA.  
 AC 009727;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonurus mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=58177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=LEMV-BRAZIL 1;  
 RA Faria J.C., Maxwell D.P.;  
 RT "Variability in geminivirus associated with Phaseolus vulgaris in  
 RT Brazil."  
 RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: U92532; AAB5157.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 FT NON\_TER 226  
 SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 78.1%; Score 282; DB 12; Length 226;  
 Best Local Similarity 78.6%; Pred. No. 4.8e-24;  
 Matches 55; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEAAVDSRGSGCOTSDAAAEALNASSKEEALQIREKIPKEYLFOFHNLNSL 60  
 DB 111 TLVWGEFQIDGSRGSGCOTANDAAAEALNASSKEEALQIREKIPKEYLFOFHNLNSL 170

QY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 4  
 09WHF6  
 ID 09WHF6 PRELIMINARY; PRT; 226 AA.  
 AC 09WHF6;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS tomato mild mottle geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
 OX NCBI\_TaxID=92943;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=HN96-HSKW;  
 RA Nakula M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
 RA Maxwell D.P.;  
 RT "Molecular characterization and DNA-based detection methods for  
 RT vegetable-infecting geminiviruses in Central America."  
 RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF131071; AAD3471.1;  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Geminl\_AL1; 1.  
 FT NON\_TER 226  
 SQ SEQUENCE 226 AA; 25941 MW; 2EA116712871A23 CRC64;

Query Match 78.1%; Score 282; DB 12; Length 226;  
 Best Local Similarity 74.3%; Pred. No. 4.8e-24;  
 Matches 52; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLVWGEAAVDSRGSGCOTSDAAAEALNASSKEEALQIREKIPKEYLFOFHNLNSL 60  
 DB 111 TLVWGEFQIDGSRGSGCOTANDAAAEALNASSKEEALQIREKIPKEYLFOFHNLNSL 170  
 QY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180

RESULT 5  
 067574  
 ID 067574 PRELIMINARY; PRT; 361 AA.  
 AC 067574;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN ALL.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 RA Maxwell D.P., Russell D.R.;  
 RT "Cloning of the complete DNA genomes of four bean-infecting  
 RT geminiviruses and determining their infectivity by electric discharge  
 RT particle acceleration."  
 RL Phytopathology 81:980-985(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 RA Morales F.J., Maxwell D.P.;  
 RT "Differentiation of bean-infecting geminiviruses by nucleic acid



RT hybridization probes and aspects of bean golden mosaic in Brazil. ";  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
 RT "Genetic diversity in geminiviruses causing bean golden mosaic  
 disease: The nucleotide sequence of the infectious cloned DNA  
 components of a Brazilian isolate of bean golden mosaic geminivirus. ";  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M88686; AAA6312.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.6%; Score 280; DB 12; Length 361;  
 Best Local Similarity 77.6%; Pred. No. 1,4e-23;  
 Matches 52; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 4 WGEAAVDRSARGCGCOTSDAAAEALNASSKEEALQITREKIPKYLFOFHNLNSLDRI 63  
 Db 113 WGFQVDRSARGCGOOTANDAAEALNASSKEEAMQITREKIPKFLFYHNLNSLDRI 172  
 QY 64 FDKTRPEP 70  
 Db 173 FTKAPDP 179

RESULT 6  
 Q90DB1 PRELIMINARY; PRT; 225 AA.  
 ID Q90DB1  
 AC Q90DB1  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=69263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CGMV-BR;  
 RA Faria J.C.;  
 RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from  
 Brazil. ";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF188708; AAF06318.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 225  
 FT SEQUENCE 225 AA; 25766 MW; 1089CB6BD8D15B5D CRC64;

Query Match 77.0%; Score 278; DB 12; Length 225;  
 Best Local Similarity 77.6%; Pred. No. 1.3e-23;  
 Matches 52; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 4 WGEAAVDRSARGCGCOTSDAAAEALNASSKEEALQITREKIPKYLFOFHNLNSLDRI 63  
 Db 113 WGFQVDRSARGCGOOTANDAAEALNASSKEEAMQITREKIPKFLFYHNLNSLDRI 172  
 QY 64 FDKTRPEP 70  
 Db 173 FTKAPDP 179

RESULT 7  
 Q9ELT8

ID Q9ELT8 PRELIMINARY; PRT; 314 AA.  
 AC Q9ELT8;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States. ";  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF288227; AAG01006.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 77.0%; Score 278; DB 12; Length 314;  
 Best Local Similarity 80.9%; Pred. No. 2e-23;  
 Matches 55; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 TIVWGEAAVDRSARGCGCOTSDAAAEALNASSKEEALQITREKIPKYLFOFHNLNSLDRI 60  
 Db 110 TITWGEVQVGRSARGCGOOTANDAAEALNASSKEEALQITREKIPKYLFOFHNLNSLDRI 169  
 QY 61 DRIFDKTP 68  
 Db 170 DRIFSPPP 177

RESULT 8  
 Q9QSS5 PRELIMINARY; PRT; 364 AA.  
 ID Q9QSS5  
 AC Q9QSS5;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States. ";  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl  
 virus : nucleotide sequence and phylogenetic relationships. ";  
 RL Online Publication.  
 DR EMBL: AF104036; AAD47173.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemini\_AL1; 1.  
 SO SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

```

Query Match          77.0%; Score 278; DB 12; Length 364;
Best Local Similarity 80.9%; Pred. No. 2.3e-23;
Matches 55; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEAAVDGRSARCGCOTSDNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSNL 60
   110 TIEWGVFOIDGRSARGGQQTANDAAAEALNSGTEKEDALKIREKLPERYLFOYHNLSSNL 169
DB 61 DRIFDKTPP 68
   170 DRIFSKPPEP 177

RESULT 9
ID 098693 PRELIMINARY; PRT; 185 AA.
AC 098693;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN AC1.
OS sida golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51034;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAMAICA;
RA Roye M.E.; McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT "Genetic diversity among geminiviruses associated with the weed
   species Sida spp, Macroplittium lathyroides, and Wissadula amplissima
   from Jamaica."
RL Plant Dis. 81:1251-1258(1997).
EMBL: U67926; AAB97865.1;
DR InterPro: IPR001191; Gemini.AL1.
DR Pfam: PF00799; Gemini.AL1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemini.AL1; 1.
FT NON_TER 185
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match          76.2%; Score 275; DB 12; Length 185;
Best Local Similarity 71.4%; Pred. No. 2.3e-23;
Matches 50; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 1 TLWGEAAVDGRSARCGCOTSDNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSNL 60
   89 TIEWGVFOIDGRSARGGQQTANDAAAEALNSGTEKEDALKIREKLPERYLFOYHNLSSNL 148
DB 61 DRIFDKTPP 70
   149 DRIFSKPPEP 158

RESULT 10
ID P88975 PRELIMINARY; PRT; 149 AA.
AC P88975;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN AC1.
OS Macroplittium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAMAICA;
RA Roye M.E.;
Thesis (1996), Biochemistry, University of the West Indies, Jamaica.

```

```

DR EMBL: U75278; AAB36919.1;
DR InterPro: IPR001191; Gemini.AL1.
DR Pfam: PF00799; Gemini.AL1; 1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemini.AL1; 1.
FT NON_TER 149
SQ SEQUENCE 149 AA; 16785 MW; E4CF5EED4C9CD508 CRC64;

Query Match          74.0%; Score 267; DB 12; Length 149;
Best Local Similarity 70.0%; Pred. No. 1.5e-22;
Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEAAVDGRSARCGCOTSDNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSNL 60
   52 TIEWGVFOIDGRSARGGQQTSDNDAAEALNSGTEKAMRVKREKLPERYLFOYHNLSSNL 111
DB 61 DRIFDKTPP 70
   112 DRIFMKDPEP 121

RESULT 11
ID 09YL44 PRELIMINARY; PRT; 233 AA.
AC 09YL44;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS Macroplittium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAMAICA STRAIN 1;
RA Roye M.E.;
RT "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
   from Jamaica."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JAMAICA STRAIN 1;
RA Roye M.E.; McLaughlin W.A., Maxwell D.P.;
RT "Molecular characterization of two distinct geminiviruses infecting M.
   lathyroides from Jamaica."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF098940; AAD17850.1;
DR InterPro: IPR001191; Gemini.AL1.
DR Pfam: PF00799; Gemini.AL1.
DR PRINTS: PR00227; GEMCOATAL1.
DR ProDom: PD000736; Gemini.AL1; 1.
FT NON_TER 233
SQ SEQUENCE 233 AA; 26356 MW; AA490AE4D2166A02 CRC64;

Query Match          74.0%; Score 267; DB 12; Length 233;
Best Local Similarity 70.0%; Pred. No. 2.5e-22;
Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEAAVDGRSARCGCOTSDNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSNL 60
   110 TIEWGVFOIDGRSARGGQQTSDNDAAEALNSGTEKAMRVKREKLPERYLFOYHNLSSNL 169
DB 61 DRIFDKTPP 70
   170 DRIFMKDPEP 179

RESULT 12
039180

```

```

Query Match:          71.5%;   Score 258;   DB 12;   Length 190;
Best Local Similarity 60.0%;   Pred. No. 2e-21;
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1
QY      1 TLWGEAANDGRARCGCCTSDMAAEALNASSKEALQITREAIPEKYLFQFINLNSL 60
      ||||| :||||||| :||| ||||||||| ||| ||||| :||| :|||:|||||||

```

RA DOI 10.1007/s12127-011-9149-7; Yamana T.  
 RT Molecular phylogeny of geminiviruses infecting wild plants in Japan.  
 RL J. Plant Res. 11:247-257(1997).  
 DR EMBL: AB001303; BAA34010.1.  
 DR J. Plant Res. 11:247-257(1997).  
 DR InterPro: IPR001191; Geminifl\_1.  
 DR Pfam: PF00709; Geminifl\_1.  
 DR PRINTS: PR00227; GEMC00TALL.  
 DR PRODOM: PD000736; Geminifl\_1.  
 DR

FT NON\_TER 1 1  
FT NON\_TER 190 190  
SQ SEQUENCE 190 AA: 21444 MW: 93C3742A8EBDB7EB CRC64;

Query Match 71.2%; Score 257; DB 12; Length 190;  
Best Local Similarity 60.0%; Pred. No. 2,6e-21;  
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

OY 1 TLVWGSAVAVDGRSARGGCGTSDNDAAAEALINASSKEEALQIREKIPEKYLFQFHNLSNL 60  
Db 85 TLEWGTFFQYDGRSARGGCGONANDACAFAALNASSKAFAALSIIREKLPKDFIFQYHNLSNL 144  
OY 61 DRI-----FDKTP 69  
Db 145 DRIFAPPLEVFVCPFSNSSSFDQVPE 169

Search completed: October 10, 2002, 02:02:00  
Job time: 1093 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:52 ; Search time 34.62 Seconds

(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346A-8

Perfect score: 361  
Sequence: 1 TLWGEAAVVDGSRAGCGCOT.....FOFHNLNSLDRIFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	350	97.0	352	1 VAL1_TGMV	P03567 tomato gold
2	250	69.3	361	1 VAL1_PYMV	P27258 potato yell
3	237	65.7	358	1 VAL1_CLVX	P14982 cassava lat
4	237	65.7	358	1 VAL1_CLVX	P14972 cassava lat
5	230	63.7	362	1 VAL1_TYLCV	P36279 tomato yell
6	223	61.8	359	1 VAL1_TYLCV	P36609 tomato yell
7	219	60.7	349	1 VAL1_PHVU	P06923 pepper huas
8	216	59.8	359	1 VAL1_TYLCV	P27260 tomato yell
9	214	59.3	353	1 VAL1_BGMV	P05175 bean golden
10	213	59.0	355	1 VAL1_ABMVW	P21947 abutilon mo
11	212	58.7	357	1 VAL1_TYLCV	P27259 tomato yell
12	210	58.2	358	1 VAL1_TYLCV	P14991 beet curly
13	210	58.2	361	1 VAL1_TYLCV	P06657 tomato yell
14	210	58.2	361	1 VAL1_TYLCV	P29048 squash leaf
15	185	57.9	347	1 VAL1_TYLCV	P16787 homo sapien
16	179	57.9	299	1 VAL1_TYLCV	P175 HELR1
17	176	57.6	447	1 VAL1_TYLCV	P175 HELR1
18	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
19	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
20	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
21	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
22	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
23	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
24	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
25	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
26	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
27	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
28	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
29	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
30	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
31	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
32	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1
33	176	57.6	1610	1 VAL1_TYLCV	P175 HELR1

34	58.5	16.2	1070	1	P11B_HUMAN	P42338 homo sapien
35	58	16.1	319	1	LDH2_SRAAM	O99435 staphylococ
36	58	16.1	617	1	YACH_ECOLI	P36682 escherichia
37	58	16.1	874	1	SLAP_BACLI	P49052 bacillus il
38	57.5	15.9	439	1	XYLA_HUMAN	O9Cf97 lactococcus
39	57.5	15.9	925	1	Y562_HUMAN	O60308 homo sapien
40	57	15.8	200	1	VIP_CHICK	P48143 gallus gall
41	57	15.8	200	1	VIP_MERGA	P45644 melaleucis g
42	57	15.8	247	1	YCP4_YEAST	P25349 saccharomyc
43	57	15.8	355	1	CRIB_RHOSH	P54905 rhodobacter
44	57	15.8	387	1	YAPF_RHOSH	P55615 rhizobium s
45	57	15.8	4639	1	DYHC_DROME	P37276 drosophila

## ALIGNMENTS

RESULT	1	STANDARD	PRT	352 AA
VAL1_TGMV	1			
ID	VAL1_TGMV			
AC	P03567			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-JUN-1994 (Rel. 29, Last annotation update)			
DE	ALI protein.			
GN	ACI.			
OS	Tomato golden mosaic virus (TGMV).			
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.			
OX	NCBI_TaxID=10831;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;			
RT	*Complete nucleotide sequence of the infectious cloned DNA components			
RT	of tomato golden mosaic virus: potential coding regions and regulatory			
RT	sequences.			
RL	EMBO J. 3:2197-2205(1984).			
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; K02029; NOT_ANNOTATED_CDS.			
DR	PIR; A04170; Q0C6V1.			
DR	InterPro; IPR001191; Gemin1_AL1.			
DR	Pfam; PF00709; Gemin1_AL1; 1.			
DR	PRINTS; PR00227; GEMCONTAL1.			
DR	ProDom; PD00736; Gemin1_AL1; 1.			
KW	ATP-binding.			
FT	NP_BIND 223			
FT	SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;			
Query Match	97.0%; Score 350; DB 1; Length 352;			
Best local Similarity	97.1%; Pred. No. 4.6e-32;			
Matches	68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
OY	1 TLWGEAAVVDGSRAGCGCOTSDAAEAALNSKKEALQIREKIPKYLEFOFHNLNSL 60			
DB	111 TLWGEAFVVDGSRAGCGCOTSDAAEAALNSKKEALQIREKIPKYLEFOFHNLNSL 170			
OY	61 DRIFDKTPEP 70			
DB	171 DRIFDKTPEP 180			
RESULT	2			
VAL1_PYMV				
ID	VAL1_PYMV			
STANDARD				
PRT	361 AA.			

```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 protein.
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311403; PubMed=1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RL potato yellow mosaic virus."
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D00940; BAA0782.1; -
DR PIR: J00364; QOCVPT.
DR InterPro: IPR001191; Geminl.AL1.
DR Pfam: PF00799; Geminl.AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Geminl.AL1; 1.
DR ATP-binding.
KW NP_BIND
FT NP_BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA; 40850 MW; 5627A3BF1264383 CRC64;

Query Match 69.3%; Score 250; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 8.4e-21;
Matches 47; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 1 TLWGFAAVDGRSARGCGTSDNDAAEALNASSKEEALQIRREKPEKYLFOFHNLNSNL 60
DB 110 TLWGQFOIDGRSARGCGQTVDAEAALNLSGTEAKAMIKREKPEKFLFOYHNLNSNL 169
OY 61 DRIEDKTPPE 69
DB 170 DRIEKKAPE 178

RESULT 3
VAL1_CLVN STANDARD; PRT; 358 AA.
ID VAL1_CLVN
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Gay M.R.;
RT "Nucleotide sequence of cassava latent virus DNA."
RL Nature 301:260-262(1983).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: J02057; -; NOT_ANNOTATED_CDS.
DR InterPro: IPR001191; Geminl.AL1.
DR Pfam: PF00799; Geminl.AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Geminl.AL1; 1.
DR ATP-binding.
KW NP_BIND
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 65.7%; Score 237; DB 1; Length 358;
Best Local Similarity 61.4%; Pred. No. 2.4e-19;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLWGFAAVDGRSARGCGTSDNDAAEALNASSKEEALQIRREKPEKYLFOFHNLNSNL 60
DB 109 TLWGQFOIDGRSARGCGQSDANDAYAKALNLSGSKSEALNYIRELVKDFVLOFHNLNSNL 168
OY 61 DRIEDKTPPE 70
DB 169 DRIEKKAPE 178

RESULT 4
VAL1_CLVN STANDARD; PRT; 358 AA.
ID VAL1_CLVN
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RL African cassava mosaic virus (Nigerian strain)."
GN Nucleic Acids Res. 18:197-198(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X17095; CAA34953.1; -
DR PIR: S07594; S07594.
DR InterPro: IPR001191; Geminl.AL1.
DR Pfam: PF00799; Geminl.AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Geminl.AL1; 1.
DR ATP-binding.
KW NP_BIND
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40435 MW; 1DB16B0CB2D5E2C CRC64;

Query Match 65.7%; Score 237; DB 1; Length 358;
Best Local Similarity 61.4%; Pred. No. 2.4e-19;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLWGFAAVDGRSARGCGTSDNDAAEALNASSKEEALQIRREKPEKYLFOFHNLNSNL 60
DB 109 TLWGQFOIDGRSARGCGQSDANDAYAKALNLSGSKSEALNYIRELVKDFVLOFHNLNSNL 168

```

OY 61 DRIEDKTEP 70  
 DB 169 DRIEOPPEP 178

# RESULT 5

VAL1\_TYLCU STANDARD: PRT: 362 AA.  
 AC P36279;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36447;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9313978; PubMed=8423446;  
 RA Dry I.B., Ridgen J.E., Krahe L.R., Mullineaux P.M., Rezaian M.A.;  
 RT "Nucleotide sequence and genome organization of tomato leaf curl  
 geminivirus";  
 RL J. Gen. Virol. 74:147-151(1993).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 DR InterPro: IPR001191; Gemini.AL1.  
 DR Pfam: PF00799; Gemini.AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRODOM: PD000736; Gemini.AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 221 228 ATP (POTENTIAL).  
 SO SEQUENCE 362 AA; 41197 MW; 343E7184B4704098 CRC64;

Query Match 63.7%; Score 230; DB 1; Length 362;  
 Best Local Similarity 54.1%; Pred. No. 1.5e-18;  
 Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

OY 1 TLVWGEAAVDRSGRCGQTSNDAAAEALNASSKEEALQITREKIPKYLFOFHINLSND 60  
 DB 110 TLWGEFOIDGRSARGGQGSANDAYAOALNKGSKSEALNVLRELAPKDYVLOFHINLSNL 169  
 OY 61 DRI-----FDKTEP 69  
 DB 170 DRIETPLEVYVSPFLSSSFDRAPE 194

# RESULT 6

VAL1\_TYLCU STANDARD: PRT: 359 AA.  
 AC P38609;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=37139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94256836; PubMed=8198442;  
 RA Norris E., Hidalgo E., Accotto G., Moriones E.;  
 RT "High similarity among the tomato yellow leaf curl virus isolates  
 from the west Mediterranean basin: the nucleotide sequence of an  
 infectious clone from Spain";  
 RL Arch. Virol. 135:165-170(1994).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: 225751; CAA81026.1; -.  
 DR PIR: S39211; S39211.  
 DR InterPro: IPR001191; Gemini.AL1.  
 DR Pfam: PF00799; Gemini.AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRODOM: PD000736; Gemini.AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 221 228 ATP (POTENTIAL).  
 SO SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 61.8%; Score 223; DB 1; Length 359;  
 Best Local Similarity 59.4%; Pred. No. 9.1e-18;  
 Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

OY 2 TLVWGEAAVDRSGRCGQTSNDAAAEALNASSKEEALQITREKIPKYLFOFHINLSND 61  
 DB 111 LFWGTFQIDGRSARGGQGTANDAYARAINNASSKEALDVYKELAPRDYILFHFNINSND 170  
 OY 62 RIFDKTEP 70  
 DB 171 RIFQVPEP 179

# RESULT 7

VAL1\_PHV STANDARD: PRT: 349 AA.  
 AC O06923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN AL1.  
 OS Pepper huasteco virus (PHV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=28349;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94015007; PubMed=8409944;  
 RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,  
 RA Rivera-Bustamante R.F.;  
 RT "Complete nucleotide sequence of pepper huasteco virus: analysis and  
 RT comparison with bipartite geminiviruses";  
 RL J. Gen. Virol. 74:2225-2231(1993).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: X70418; CAA49856.1; -.  
 DR PIR: S31875; S31875.  
 DR PIR: J02300; J02300.  
 DR InterPro: IPR001191; Gemini.AL1.  
 DR Pfam: PF00799; Gemini.AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRODOM: PD000736; Gemini.AL1.1.  
 KW ATP-binding.  
 FT NP\_BIND 221 228 ATP (BY SIMILARITY).  
 SO SEQUENCE 349 AA; 39722 MW; D5F4E76CD56370F4 CRC64;

Query Match 60.7%; Score 219; DB 1; Length 349;





```

DR EMBL: X15983; -: NOT_ANNOTATED_CDS.
DR PIR: A36214; QOCVWL.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.
DR ATP-binding.
KW NP_BIND
FT SEQUENCE 355 AA; 40678 MW; 16A2CA8A63251E95 CRC64;
SQ

Query Match
Best Local Similarity 59.0%; Score 213; DB 1; Length 355;
Matches 40; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

OY 1 TLWGEAAVDCRSARGCGTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSNL 60
DB 110 TAEWGEFQIDGRSARGCGQTANDSYAKALNAGVQSAALNLIKKEPKDYVLQHNHNSNL 169
OY 61 DRIFDKTRPEP 70
DB 170 ERIFFAKRPEP 179

RESULT 11
VAL1_TYLCV STANDARD; PRT; 357 AA.
ID VAL1_TYLCV
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (C1 protein).
GN C1.
OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024070; PubMed=1926771;
RA Navot N., Pichersky E., Zeldan M., Zamir D., Czosnek H.;
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT with a single genomic component."
RL Virology 185:151-161(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15656; CAA33688.1; -.
DR PIR: D40779; QOCVCI.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.
DR ATP-binding.
KW NP_BIND
FT SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;
SQ

Query Match
Best Local Similarity 58.7%; Score 212; DB 1; Length 357;
Matches 40; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

OY 4 WGEAAVDCRSARGCGTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSLDRI 63
DB 111 FGVSQIDRSARGGQGSANDAYAEALNAGSSEALNLIKKEAPKDYVLQFHNLSSNLDRI 170
OY 64 F 64

```

```

DB 171 F 171

RESULT 12
VAL1_BCTV STANDARD; PRT; 358 AA.
ID VAL1_BCTV
AC P14991;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 protein (40.8 kDa protein).
OS Beet curly top virus (BCTV).
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX NCBI_TaxID=10840;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT "The nucleotide sequence of an infectious clone of the geminivirus
RT beet curly top virus."
RL EMBO J. 5:1761-1767(1986).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04144; -: NOT_ANNOTATED_CDS.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.
DR ATP-binding.
KW NP_BIND
FT SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C33 CRC64;
SQ

Query Match
Best Local Similarity 58.2%; Score 210; DB 1; Length 358;
Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLWGEAAVDCRSARGCGTSDNDAAEALNASSKEEALQITREKIPKYLFOFHNLNSNL 60
DB 110 TIWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEPKDYFLQHHNLNNA 169
OY 61 DRIFDKTRPEP 70
DB 170 QKIFQRPDP 179

RESULT 13
VAL1_TMOV STANDARD; PRT; 361 AA.
ID VAL1_TMOV
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AL1.
OS Tomato mottle virus (isolate Florida) (TMOV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=36449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107858; PubMed=1469361;
RA Abouzid A.M., Polston J.E., Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT isolated from tomatoes in Florida."
RL J. Gen. Virol. 73:3225-3229(1992).

```

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L14460; AAC32414.1; -  
CC PIR: J01870; J01870.  
CC InterPro: IPR001191; Gemin1.AL1.  
CC Pfam: PF00799; Gemin1.AL1; 1.  
CC PRINTS: PR00227; GEMCOATL1.  
CC ProDom: PD000736; Gemin1.AL1; 1.  
CC ATP-binding.  
CC NP\_BIND 222 229 ATP (BY SIMILARITY).  
CC SEQUENCE 361 AA; 40516 MW; 813865CEAC6950 CRC64;  
CC  
CC Query Match 58.2%; Score 210; DB 1; Length 361;  
CC Best Local Similarity 54.3%; Pred. No. 2.7e-16;  
CC Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;  
CC  
CC QY 1 TLVWGAADGRRAGGCGTSDNDAAEALNASSKEEALQIREKIPEKYLFPFHNLNSNL 60  
CC ID 110 TIEMGFQJIDGRRAGGGSANDSYAKALNAGSVSALVALKEQPKDFVIONHNIRSL 169  
CC Db 61 DRIFDKTPEP 70  
CC :||| |||  
CC 170 ERIFAKAPEP 179  
CC  
CC RESULT 14  
CC VAL1\_SLCV STANDARD; PRT; 347 AA.  
CC ID VAL1\_SLCV  
CC AC P29048;  
CC DT 01-DEC-1992 (Rel. 24, Created)  
CC DT 01-DEC-1992 (Rel. 24, Last sequence update)  
CC DT 01-DEC-1992 (Rel. 24, Last annotation update)  
CC DE AL1 protein.  
CC OS Squash leaf curl virus.  
CC OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
CC OX NCBI\_TaxID=10829;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RA MEDLINE=91082449; PubMed=1984668;  
CC RA Lazarowitz S.G.; Lazdins J.B.;  
CC RT Infectivity and complete nucleotide sequence of the cloned genomic  
CC RT components of a bipartite squash leaf curl geminivirus with a broad  
CC RT host range phenotype.";  
CC RL Virology 180:58-69(1991).  
CC  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M38183; AAC32410.1; ALT\_INIT.  
CC PIR: C36785; QOCYS1.  
CC InterPro: IPR001191; Gemin1.AL1.  
CC Pfam: PF00799; Gemin1.AL1; 1.  
CC PRINTS: PR00227; GEMCOATL1.  
CC ProDom: PD000736; Gemin1.AL1; 1.  
CC ATP-binding.  
CC NP\_BIND 218 225 ATP (POTENTIAL).  
CC SEQUENCE 347 AA; 39110 MW; AFDABDEE122110E CRC64;  
CC

CC Query Match 32.4%; Score 117; DB 1; Length 347;  
CC Best Local Similarity 37.9%; Pred. No. 7.3e-06;  
CC Matches 25; Conservative 12; Mismatches 25; Indels 4; Gaps 1;  
CC  
CC QY 5 GEAAVDGSRAGGCGTSDNDAAEALNASSKEEALQIREKIPEKYLFPFHNLNSNLDRIF 64  
CC ID 116 GQRYVSG-----GSKSNKDDYHNAVNAGSAGEALDIKADPKTFIVNHLANVERLF 171  
CC Db 65 DKTPEP 70  
CC :||| |||  
CC 172 QKPEP 177  
CC  
CC RESULT 15  
CC LMA3\_HUMAN STANDARD; PRT; 1713 AA.  
CC ID LMA3\_HUMAN  
CC AC Q16787; Q13679; Q13680;  
CC DT 01-NOV-1997 (Rel. 35, Created)  
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)  
CC DT 01-MAR-2002 (Rel. 41, Last annotation update)  
CC DE Laminin alpha-3 chain precursor (Epiligrin 170 kDa subunit) (E170)  
CC DE (Nuclein alpha subunit).  
CC LAMA3.  
CC OS Homo sapiens (Human).  
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
CC OX NCBI\_TaxID=9606;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RC TISSUE=Keratinocytes;  
CC RX MEDLINE=94357926; PubMed=8077230;  
CC RA Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;  
CC RT "Cloning of the Lama3 gene encoding the alpha 3 chain of the adhesive  
CC RT ligand epiligrin. Expression in wound repair.";  
CC RL J. Biol. Chem. 269:22779-22787(1994).  
CC  
CC [2]  
CC RP SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).  
CC RX MEDLINE=96163880; PubMed=8586427;  
CC RA Vidal F., Baudoin C., Miquel C., Galliano M.-F., Cristiano A.M.,  
CC RA Uitto J., Ortonne J.-P., Meneguzzi G.;  
CC RT "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification  
CC RT of a homozygous deletion in a patient with Herlitz junctional  
CC RT epidermolysis bullosa.";  
CC RL Genomics 30:273-280(1995).  
CC  
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
CC  
CC -1- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION  
CC VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-  
CC 6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE  
CC PHOSPHORYLATION OF P125-FAK AND P80, (3) DIFFERENTIATION OF  
CC KERATINOCYTES.  
CC  
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
CC THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIRIN/KALININ/  
CC NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)  
CC AND LAMININ-7 (KS-LAMININ).  
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT  
CC MEMBRANES (MAJOR COMPONENT).  
CC  
CC -1- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-  
CC TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT  
CC A.  
CC  
CC -1- TISSUE SPECIFICITY: SKIN; RESPIRATORY; URINARY; AND DIGESTIVE  
CC EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT  
CC SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,  
CC AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A  
CC DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED  
CC IN THE CNS.  
CC  
CC -1- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.  
CC

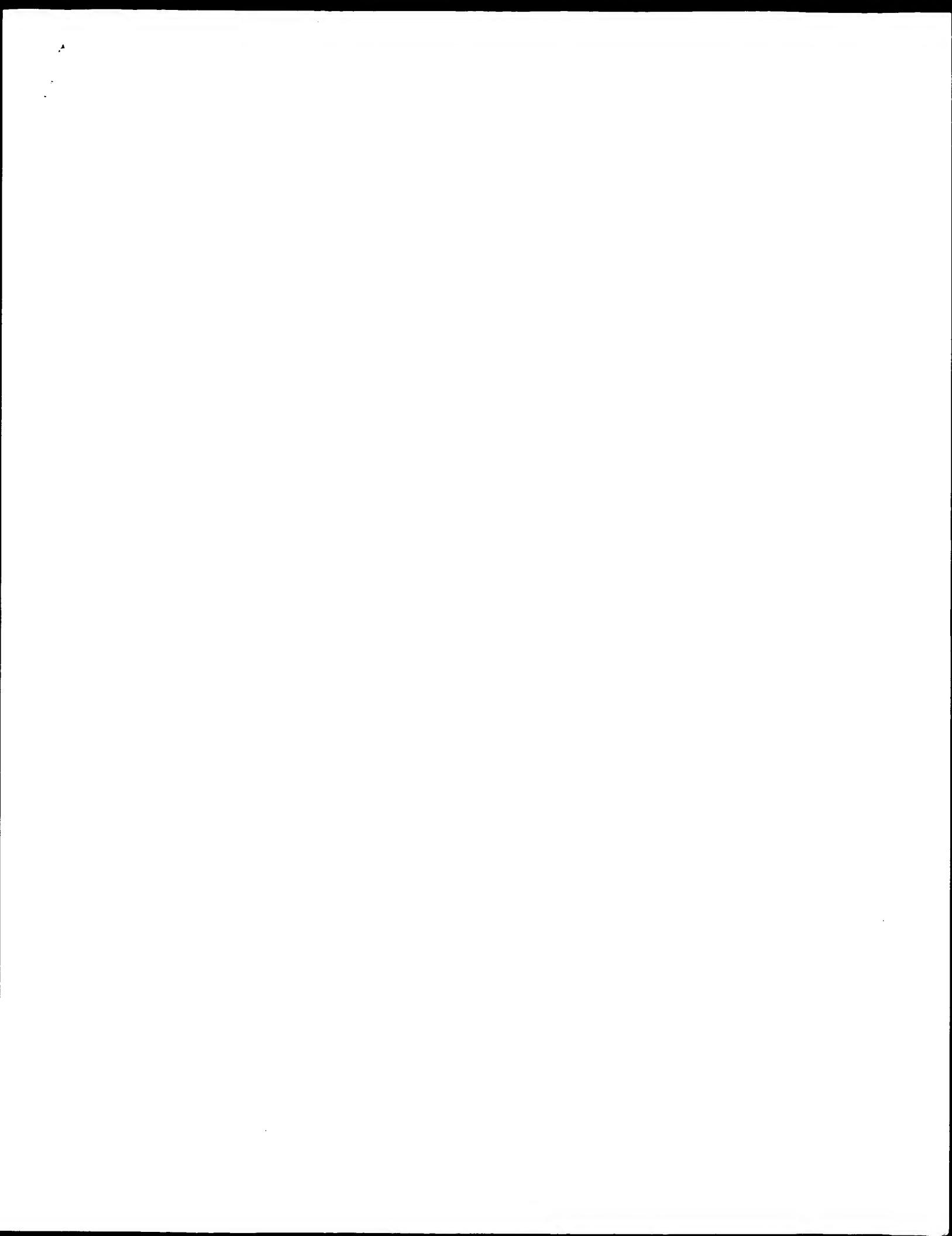
```

FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1108 1108 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1131 1131 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1325 1325 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1477 1477 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1667 1667 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1 45 MGMTIFGALGCLGTSOQKRVPELQPPGOSLOASTIVE
FT FRPS -> KVSSTGYGLTYQAKSFALGDMVLLEKRPVQL
FT TGOHMSIYEETNTPRPDRLHHGRVHVVEGNEFRHASSRAPV
FT SREELMTVLSRLADVRIOGLYFETETORLTLSVEGLEASDT
FT GSGRIALAVEICACPPAYAGDSC (IN ISOFORM B).
FT FT CONFLICT 5 5 ATG -> GMC (IN REF. 2).
FT FT CONFLICT 123 125 M -> K (IN REF. 2).
FT FT CONFLICT 481 481 M -> L (IN REF. 2).
FT FT CONFLICT 754 754 R -> Q (IN REF. 2).
FT FT CONFLICT 969 969 E -> A (IN REF. 2).
FT FT CONFLICT 1052 1052 D -> A (IN REF. 2).
FT FT CONFLICT 1184 1184 G -> A (IN REF. 2).
SQ SEQUENCE 1713 AA; 189304 MW; 45EA9BE1017B60D3 CRC64;

Query Match 19.0%; Score 68.5; DB 1; Length 1713;
Best Local Similarity 35.5%; Pred No.13;
Matches 22; Conservative 12; Mismatches 21; Indels 7; Gaps 3
QY 8 AYDGRSARGGCGOTNSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLSNLDRIIF 64
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 552 AYDAATATYENILNAIKAAEDDANBRAASASESALQTVIKEDLPK-----AKTLSSNSDKLL 607
QY 65 DK 66
   : :
Db 608 NE 609

Search completed: October 10, 2002, 02:09:53
Job time: 1351 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:45 ; Search time 69.63 seconds

(without alignments)  
96.600 Million cell updates/sec

Title: US-09-289-346a-8

Perfect score: 361

Sequence: 1 TLVWGEAAVDGRSARSGCQT.....FOFHNLSNLDRIPTKPEP 70

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR 71:\*

1: PIR1:1  
2: PIR2:1  
3: PIR3:1  
4: PIR4:1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	350	97.0	352	1 QOCVLI	ALI protein - toma
2	250	69.3	361	1 QOCVPT	ALI protein - toma
3	237	65.7	358	2 S07594	hypothetical prote
4	230	63.7	362	1 J01887	ALI protein - toma
5	223	61.8	359	2 S39211	gene C1 protein -
6	219	60.7	349	2 J02300	replisase - pepper
7	216	60.7	349	2 S31875	ALI protein - pep
8	216	59.8	359	2 S22593	hypothetical prote
9	213	59.0	351	2 J02327	ALI protein - indi
10	213	59.0	355	1 QOCVWL	AVI protein - abut
11	212	58.7	357	1 QOCVCI	ALI protein - toma
12	210	58.2	358	1 J01870	gene C1 protein -
13	210	58.2	359	2 S39225	replication-associ
14	210	58.2	385	2 S28360	ALI protein - beet
15	203	56.2	360	2 S59885	ACI protein (clone
16	142	39.3	131	2 S45059	adhesive ligand ep
17	117	32.4	347	1 QOCVSI	Fe gamma (196) rec
18	68.5	19.0	1713	2 A55347	hypothetical prote
19	68	18.8	587	2 J01419	probable flagellar
20	67	18.6	180	2 D84082	hypothetical prote
21	66.5	18.4	201	2 A81380	probable peptidyl-
22	64.5	17.9	299	2 T12544	voltage-dependent
23	63.5	17.6	447	2 B71967	voltage-dependent
24	63.5	17.6	1646	2 JH0422	calcium channel at
25	63.5	17.6	2161	2 JH0564	calcium channel at
26	63.5	17.6	2161	2 A38198	voltage-dependent
27	63.5	17.6	2161	2 T42742	nitrogen regulator
28	63.5	17.6	2203	2 AD3475	
29	62	17.2	154	2	

30	62	17.2	181	2 G97976	conserved hypothet
31	62	17.2	295	2 D42452	C1 protein - tobac
32	62	17.2	714	2 C95382	probable ferricpro
33	61.5	17.0	481	2 A70091	probable phosphes
34	61	16.9	1502	1 RGBYH1	CY1/CY3 transcri
35	60	16.6	432	2 D75348	gamma-glutamyl pho
36	60	16.6	447	2 S52437	CDP-diacylglycerol
37	59.5	16.5	136	2 T22240	38k ribosome-assoc
38	59.5	16.5	269	2 J07700	cell binding facto
39	59.5	16.5	299	2 G64541	hypothetical prote
40	59.5	16.5	388	2 C69196	topoisomerase iv s
41	59.5	16.5	642	2 D90558	hypothetical prote
42	59.5	16.5	1033	2 B97700	hypothetical prote
43	59.5	16.5	2137	2 T05244	conserved hypothet
44	59	16.3	316	2 C82085	probable NADH oxid
45	59	16.3	397	2 B71078	

## ALIGNMENTS

### RESULT 1

QOCVLI ALI protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence.revision 28-Aug-1985 #text-change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stehn, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus ALI protein

Query Match

Best Local Similarity 97.1%: Score 350; DB 1; Length 352;

Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TLVWGEAAVDGRSARSGCQTSNDAAEALNASSKEALQITREKIPKYLFOFHNLSNL 60

DB 111 TLVWGEFQVDGRSARSGCQTSNDAAEALNASSKEALQITREKIPKYLFOFHNLSNL 170

QY 61 DRIFDKTPEP 70

DB 171 DRIFDKTPEP 180

RESULT 2

QOCVPT ALI protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence.revision 30-Jun-1992 #text-change 16-Jun-2000

C:Accession: J00364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

RESULT	7
S31875	
ALL protein	- pepper rizado amarillo virus
C;Species:	pepper rizado amarillo virus







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:19 : Search time 133.15 Seconds  
(without alignments)  
58.394 Million cell updates/sec

Title: US-09-289-346a-8  
Perfect score: 361  
Sequence: 1 TLVWGFRAAVDGRSARGCQOT.....FQPHNINSMLDRIFDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	361	100.0	70	21	AA18684
2	350	97.0	70	21	AA18677
3	350	97.0	356	21	AA18687
4	342	94.7	70	21	AA18685
5	339	93.9	70	21	AA18688
6	338	93.6	70	21	AA18692
7	336	93.1	70	21	AA18690
8	335	92.8	70	21	AA18678
9	335	92.8	70	21	AA18686
10	334	92.5	70	21	AA18689
11	332	92.0	70	21	AA18680

12	332	92.0	70	21	AA18691	Mutant peptide der
13	329	91.1	70	21	AA18681	Mutant peptide der
14	328	90.9	70	21	AA18683	Mutant peptide der
15	326	90.3	70	21	AA18682	Mutant peptide der
16	318	88.1	70	21	AA18679	Mutant peptide der
17	316	88.1	70	21	AA18638	Mutant peptide der
18	316	88.1	70	21	AA18638	Mutant peptide der
19	316	88.1	70	21	AA18638	Mutant peptide der
20	316	88.1	70	21	AA18638	Mutant peptide der
21	316	88.1	70	21	AA18638	Mutant peptide der
22	316	88.1	70	21	AA18638	Mutant peptide der
23	316	88.1	70	21	AA18638	Mutant peptide der
24	316	88.1	70	21	AA18638	Mutant peptide der
25	316	88.1	70	21	AA18638	Mutant peptide der
26	316	88.1	70	21	AA18638	Mutant peptide der
27	316	88.1	70	21	AA18638	Mutant peptide der
28	316	88.1	70	21	AA18638	Mutant peptide der
29	316	88.1	70	21	AA18638	Mutant peptide der
30	316	88.1	70	21	AA18638	Mutant peptide der
31	316	88.1	70	21	AA18638	Mutant peptide der
32	316	88.1	70	21	AA18638	Mutant peptide der
33	316	88.1	70	21	AA18638	Mutant peptide der
34	316	88.1	70	21	AA18638	Mutant peptide der
35	316	88.1	70	21	AA18638	Mutant peptide der
36	316	88.1	70	21	AA18638	Mutant peptide der
37	316	88.1	70	21	AA18638	Mutant peptide der
38	316	88.1	70	21	AA18638	Mutant peptide der
39	316	88.1	70	21	AA18638	Mutant peptide der
40	316	88.1	70	21	AA18638	Mutant peptide der
41	316	88.1	70	21	AA18638	Mutant peptide der
42	316	88.1	70	21	AA18638	Mutant peptide der
43	316	88.1	70	21	AA18638	Mutant peptide der
44	316	88.1	70	21	AA18638	Mutant peptide der
45	316	88.1	70	21	AA18638	Mutant peptide der

## ALIGNMENTS

RESULT 1	AA18684	standard; peptide; 70 AA.
ID	AA18684	
AC	AA18684	
XX	22-JAN-2001	(first entry)
DT		
XX		
DE		Mutant peptide derived from amino acids 110-179 of Rep (A11) protein.
XX		
KW		Geminivirus; replication protein; Rep protein; A11; transgenic plant; ribosome binding region; resistance; geminivirus infection.
XX		
OS		Synthetic.
OS		Tomato golden mosaic virus.
FX		Key
FT		Misc-difference 7 Location/Qualifiers
FT		Misc-difference 8 /note="wild type residue replaced with Ala"
FT		
PN		MO200054573-A1.
XX		
PD		21-SEP-2000.
XX		
PF		15-MAR-2000; 2000WO-US06759.
XX		
PR		18-MAR-1999; 99US-0125004.
PR		09-APR-1999; 99US-0289346.
XX		(UNCL-) UNIV NORTH CAROLINA STATE.
PA		
XX		
PI		Hanley-Bowdoin L., Orozco BM., Kong L.

XX WPI: 2000-618851/59.  
 DR Transgenic plants with increased resistance to geminivirus infection  
 XX comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant ALI protein with a mutation in the Rb binding region  
 PT -  
 PS Claim 52; Page 45; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as ALI. ALI binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the ALI  
 CC protein are used to produce transgenic plants. The mutation in ALI is  
 CC present in a ribosome binding region, and expression of mutant ALI  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant ALI proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;

Query Match 100.0%; Score 361; DB 21; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-38;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLWGEAAVDSRGARCCGCTSDNAAAEALNASSKEEALQIRREKIPKYLFPFHNLNSL 60  
 Db 1 tlwgeaavdgsrsgcgtlsndaaaealnasskeaalqirrekipekylfghnlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 2  
 AAB18677  
 ID AAB18677 standard; peptide; 70 AA.  
 AC AAB18677;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 XX  
 KW Geminivirus; replication protein; Rep protein; ALI; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI: 2000-618851/59.  
 XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant ALI protein with a mutation in the Rb binding region

PT encoding a mutant ALI protein with a mutation in the Rb binding region  
 PT -  
 PS Disclosure; Page 18; 73pp; English.  
 XX  
 XX The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as ALI. ALI binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the ALI protein are used  
 CC to produce transgenic plants. The mutation in ALI is present in a  
 CC ribosome binding region, and expression of mutant ALI protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant ALI  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX  
 SQ Sequence 70 AA;

Query Match 97.0%; Score 350; DB 21; Length 70;  
 Best Local Similarity 97.1%; Pred. No. 1.5e-36;  
 Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TLWGEAAVDSRGARCCGCTSDNAAAEALNASSKEEALQIRREKIPKYLFPFHNLNSL 60  
 Db 1 tlwgeaavdgsrsgcgtlsndaaaealnasskeaalqirrekipekylfghnlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 3  
 AAB18687  
 ID AAB18687 standard; peptide; 356 AA.  
 AC AAB18687;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Amino acid sequence of a geminivirus replication protein of TGMV.  
 XX  
 KW Geminivirus; replication protein; Rep protein; ALI; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key location/Qualifiers  
 FT Misc-difference 356  
 FT /note= "unspecified amino acid"

WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI: 2000-618851/59.  
 XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant ALI protein with a mutation in the Rb binding region

PT -  
XX  
PS Disclosure: Page 47-48: 73pp: English.  
XX

CC The present sequence represents a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
SQ Sequence 356 AA;  
XX

Query Match 97.0%; Score 350; DB 21; Length 356;  
Best Local Similarity 97.1%; Pred. No. 1.2e-35;  
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TLWGEAAVDGRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
Db 110 TLWGEfgyvdgrsargcgctsdnaaaalnasskeeaqlirekipekylfqfhnlnsl 169  
QY 61 DRIFDKTPEP 70  
Db 170 drifdktp 179

## RESULT 4

AAB18685 AAB18685 standard; peptide; 70 AA.

AC AAB18685;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
OS Tomato golden mosaic virus.

OS Synthetic.  
FH Key Location/Qualifiers

FT Misc-difference 10 /note= "wild type residue replaced with Ala"

FT MO200054573-A1.

PD 21-SEP-2000.

PD 15-MAR-2000; 2000MO-US06759.

PF 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

PA Hanley-Bowdoin L, Orozco BM, Kong L;

PI WPI; 2000-618851/59.

DR Transgenic plants with increased resistance to geminivirus infection  
XX comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region

PT -  
XX  
PS Claim 53; Page 46: 73pp: English.  
XX

CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
SQ Sequence 70 AA;  
XX

Query Match 94.7%; Score 342; DB 21; Length 70;  
Best Local Similarity 95.7%; Pred. No. 1.5e-35;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLWGEAAVDGRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
Db 1 TLWGEfgyvdgrsargcgctsdnaaaalnasskeeaqlirekipekylfqfhnlnsl 169  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

## RESULT 5

AAB18688 AAB18688 standard; peptide; 70 AA.

AC AAB18688;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
OS Tomato golden mosaic virus.

OS Synthetic.  
FH Key Location/Qualifiers

FT Misc-difference 19 /note= "wild type residue replaced with Ala"

FT Misc-difference 20 /note= "wild type residue replaced with Ala"

FT MO200054573-A1.

PD 21-SEP-2000.

PD 15-MAR-2000; 2000MO-US06759.

PF 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

PA Hanley-Bowdoin L, Orozco BM, Kong L;

PI WPI; 2000-618851/59.

DR Transgenic plants with increased resistance to geminivirus infection  
XX comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region



```
XX WPI; 2000-618651/59.
DR Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
PS Disclosure; Page 49; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 93.1%; Score 336; DB 21; Length 70;
Best Local Similarity 94.3%; Pred. No. 8.7e-35;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TLVWGEAAVDSRGSGCOTSDNDAEAALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60
Db 1 tlwgefvdgsgsgcotsndaaalaasskeeaqlirekipekylfqfhnlnsl 60
OY 61 DRIFDKTPEP 70
Db 61 drifdktp 70

RESULT 8
AAB18678
ID AAB18678 standard; peptide; 70 AA.
XX
AC AAB18678;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
XX
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 12 /note= "wild type residue replaced with Ala"
FT Misc-difference 13 /note= "wild type residue replaced with Ala"
FT Misc-difference 15 /note= "wild type residue replaced with Ala"
FT /note= "wild type residue replaced with Ala"
XX
XX WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000MO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX
XX 09-APR-1999; 99US-0289346.
XX
```

```
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI; 2000-618651/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
PS Claim 53; Page 42-43; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 92.8%; Score 335; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 1.2e-34;
Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 TLVWGEAAVDSRGSGCOTSDNDAEAALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60
Db 1 tlwgefvdgsgsgcotsndaaalaasskeeaqlirekipekylfqfhnlnsl 60
OY 61 DRIFDKTPEP 70
Db 61 drifdktp 70
```

```
RESULT 9
AAB18686
ID AAB18686 standard; peptide; 70 AA.
XX
AC AAB18686;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
XX
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 24 /note= "wild type residue replaced with Leu"
FT Misc-difference 25 /note= "wild type residue replaced with Leu"
FT Misc-difference 26 /note= "wild type residue replaced with Leu"
FT /note= "wild type residue replaced with Leu"
XX
XX WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX
```

PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 53; Page 46; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:  
 Query Match 92.8%; Score 335; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 1.2e-34;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLVWGSAVVDGRSARGCGCTSDNDAAEALNASSKEEALQIRKIPKYLFGHINSL 60  
 Db 1 tlwvgefydvgrsargcgctsdnlllealnasskeeaqlirekipekylfgfhlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70  
 Db 61 drifdktp 70  
 RESULT 10  
 AAB18689  
 ID AAB18689 standard; peptide; 70 AA.  
 XX  
 AC AAB18689;  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 22 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 23 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 XX

PD 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure; Page 48-49; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:  
 Query Match 92.5%; Score 334; DB 21; Length 70;  
 Best Local Similarity 94.3%; Pred. No. 1.6e-34;  
 Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 TLVWGSAVVDGRSARGCGCTSDNDAAEALNASSKEEALQIRKIPKYLFGHINSL 60  
 Db 1 tlwvgefydvgrsargcgctsaanaealnasskeeaqlirekipekylfgfhlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70  
 Db 61 drifdktp 70  
 RESULT 11  
 AAB18680  
 ID AAB18680 standard; peptide; 70 AA.  
 XX  
 AC AAB18680;  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 42 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 43 /note= "wild type residue replaced with Ala"  
 FT  
 XX  
 FT Misc-difference 44

FT /note= "wild type residue replaced with Ala"  
XX  
XX WO200054573-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX WPI: 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
PS Claim 52; Page 43-44; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX Sequence 70 AA:  
SQ  
Query Match 92.0%; Score 332; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 2.8e-34;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLWGEEAVDGRSARGCGCOTSDNDAAEALNASSKEEALQIRREKIPKYLFOFHNLNSNL 60  
Db 1 tlwgefgvqgrsarqgcqtlnsdadaaalnasskeaalqliaaapekylfqnlnsnl 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70  
RESULT 12  
AAB18691  
ID AAB18691 standard; peptide; 70 AA.  
XX  
XX AAB18691;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
XX  
XX Tomato golden mosaic virus.  
XX  
XX Key location/Qualifiers  
FH Misc-difference 34  
FT

FT /note= "wild type residue replaced with Ala"  
FT Misc-difference 35  
FT /note= "wild type residue replaced with Ala"  
FT Misc-difference 36  
FT /note= "wild type residue replaced with Ala"  
XX  
XX WO200054573-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX WPI: 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
PS Disclosure; Page 49; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX Sequence 70 AA:  
SQ  
Query Match 92.0%; Score 332; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 2.8e-34;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLWGEEAVDGRSARGCGCOTSDNDAAEALNASSKEEALQIRREKIPKYLFOFHNLNSNL 60  
Db 1 tlwgefgvqgrsarqgcqtlnsdadaaalnasskeaalqliaaapekylfqnlnsnl 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70  
RESULT 13  
AAB18681  
ID AAB18681 standard; peptide; 70 AA.  
XX  
XX AAB18681;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
OS

[illegible]

KW	Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
RN	ribosome binding region; resistance; geminivirus infection.
OS	Synthetic.
XX	Tomato golden mosaic virus.
XX	
FT	Key Location/Qualifiers
FF	Misc-difference 59 /note= "wild type residue replaced with Ala"
FT	Misc-difference 61 /note= "wild type residue replaced with Ala"
FT	Misc-difference 62 /note= "wild type residue replaced with Ala"
XX	w0200054573-A1.
PX	21-SEP-2000.
XX	
PF	15-MAR-2000; 2000MO-US06759.
PR	18-MAR-1999; 99US-0125004.
PR	09-APR-1999; 99US-0289346.
PA	(UNCL) UNIV NORTH CAROLINA STATE.
XX	Hanley-Bowdoin L, Orozco BM, Kong L;
XI	WPI; 2000-618851/59.
DR	
XX	
PT	Transgenic plants with increased resistance to geminivirus infection
PT	comprise a nucleic acid construct containing a nucleic acid sequence
PT	encoding a mutant A1 protein with a mutation in the Rb binding region
PS	-
XX	
Claim 53; Page 45; 73pp: English.	
CC	The present sequence represents a mutant peptide, derived from a
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC	double-stranded DNA, catalyzes cleavage and ligation of single-stranded
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC	protein are used to produce transgenic plants. The mutation in AL1 is
CC	present in a ribosome binding region, and expression of mutant AL1
CC	protein imparts increased resistance to geminivirus infection in the
CC	plant. Mutant AL1 proteins are useful for producing plants having
CC	increased resistance or reduced sensitivity to a geminivirus such as
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC	virus, cotton leaf curl virus or beet curly top virus.
XX	
SQ	Sequence 70 AA:
Query Match	90.9%; Score 328; DB 21; Length 70;
Best Local Similarity	92.9%; Pred. No. 8,9e-34;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY 1 TLVGEAANDGRSARGCGCTSDAAEAALNMSSKEEALQLITREKIPKYLFQGHINLSNL 60	
Db 1 tlvgefgvdygrsarqcgctsdnaaaalnasskeaalirekipekyilfghinsal 60	
QY 61 DRIEDKTPEP 70	
Db 61 aaIdktpep 70	
RESULT 15	
AAB18682	
ID AAB18682 standard; peptide; 70 AA.	
XX	
XX	
XX	
XX	
XX	

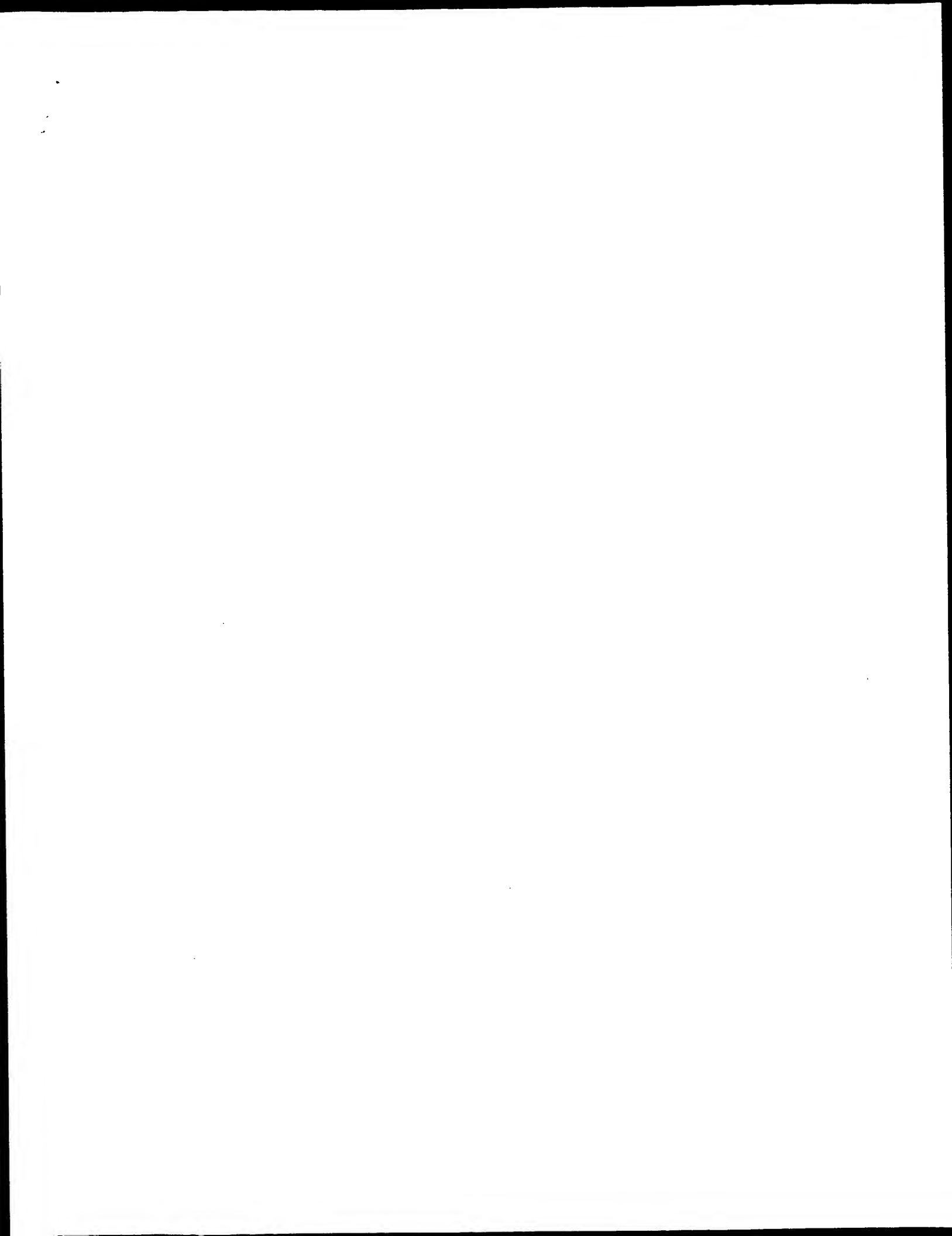


Job time: 528 sec

DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 52 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 54 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
 FT  
 XX /note= "wild type residue replaced with Ala"  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPL; 2000-618851/59.  
 DR  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 53; Page 44-45; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:

Query Match 90.3%; Score 326; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 1.6e-33;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 TLVWGEAAVDSRARGCGQTNDAAAEALNMSKEEALQITREKIPKYLFOFHNLSNL 60  
 Db 1 TLVWGEAAVDSRARGCGQTNDAAAEALNMSKEEALQITREKIPKYLFOFHNLSNL 60  
 OY 61 DRIFDKTPEP 70  
 Db 61 DRIFDKTPEP 70

Search completed: October 10, 2002, 01:47:19





AC Q9E000: 2001 (TReMBLrel\_16, Created)  
DT 01-MAR-2001 (TReMBLrel\_16, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel\_16, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel\_17, Last annotation update)  
DE REPLICATION-ASSOCIATED PROTEIN.  
GN AC1.  
OS Tomato rugose mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=134599;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.C.,  
RT Zamboli E.M., Zerbini F.M.:  
RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
RT (TRMV), a Begomovirus Isolated from Tomatoes at Triângulo Mineiro,  
RT Minas Gerais, Brazil."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF291705; AAC15546.1; -;  
DR InterPro: IPR001191; GeminiAL1.  
DR Pfam: PF00799; Gemini\_AL1.1.  
DR PRINTS: PR00227; GEMCOTALL.  
DR PRODOM: PD000736; Gemini\_AL1.1.  
SQ SEQUENCE 352 AA; 40012 MW; 47CD55838E24D613 CRC64;

Query Match	84.3%;	Score 305;	DB 12;	Length 352;
Best Local Similarity	-81.4%;	Pred. No. 8.2e-27;		
Matches 57; Conservative	7;	Mismatches 6;	Indels 0;	Gaps 0;

```
QY      1 TLVGEFQVAGRSARSGCQTSDAAAEALNASSKEEALOIRREKPEKYLFQFHNLNSL 60
        I : ::::: :::::::::::::::::::: ::::::::::::::: :::::::::::::::
Db     111 TIENGEEQLDGRSARGCGQTANDAAAEALNAPSKDIALQIRREKPEKFLFQFHNLNSL 170
```

QY	61	DRI	FKT	PEP	70
			:		
Db	171	DRIF	ARAP	EP	180

009727	PRELIMINARY;	PRT;	226 AA
ID 009727			

DT	01-JUL-1997	(TREMblrel_04, Created)
DT	01-JUL-1997	(TREMblrel_04, last sequence update)
DT	01-DEC-2001	(TREMblrel_19, last annotation update)
DE	REP PROTEIN (FRAGMENT)	

OC *Leontolus mosaii* virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus  
OX NCBI\_TaxID=58177;

RP	SEQUENCE FROM N. A.
RC	STRATIN=LEMV- BRAZIL

RT "Variability in geminivirus associated with Phaseolus vulgaris in

Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases

DR InterPro; IPR001191; Gemini\_ALL.

DR PRINTS; PR00227; GEMCOATALL.  
DR Prodom; PD000736; Gemini ALL;

NOV_18_20	NOV_18_20
SEQUENCE	226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;
SQ	

Query Match	79.6%;	Score 288;	DB 12;	Length 226;
Best Local Similarity	80.0%;	Pred. No. 4.3e-25;		
Matches	56;	Conservative	4;	Mismatches 10;
			Indels	0;
			Gaps	0

QY 1 TLWGEFVOAGRSARGCCQIISNDAAALNASSKEEALOIREKLIPEKYLFQHNLNSL 60  
| : ||||| | : ||||| | : ||||| | : ||||| | : ||||| | : ||||| | : |||||  
Db 111 TVEGGEFFQVGRSSRGGQQTVDAAAEALNPDKRTALOIKEKLEPKYLQFHNLNSL 170

QY	61	DRIE	DKT	PEP	70
Db	171	DRIE	AKA	PEP	180

RESULT 4  
Q9WHF6  
ID Q9WHF6 PRELIMINARY;  
PRT; 226 AA

DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-JUN-2001 (TREMBlere). 17, Last annotation update)  
DE REPLICATION-ASSOCIATED PROTEIN (EPACMNM)

GN REP.  
GN + COM + C E f j 7 EC + j 0 conf d f r r r u s

OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
OX NCBI TaxID=92943.

RN	[1]	SEQUENCE FROM N.A.
RD		

RC STRAIN=HN96-H5KW;  
RA Nekhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.

RT "Molecular characterization and DNA-based detection methods for RA Maxwell D.P.;

Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases

DR EMBL; AF131071; AAD33471.1;  
DR InterPro; IPR001191; Gemini\_4

```
DR      ram; pf00/99; Gemini_ALL; 1.  
DR      PRINTS; PR00227; GEMCOATALL.
```

DR	1	PROB01; PD000738; GEMINI_ALL; 1
FT	NON_TER	226 226

SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match	79.68;	Score 288;	DB 12;	Length 226;
Best Local Similarity	75.78;	Pred. No. 4.3e-25;		
Matches	53;	Conservative	12;	Mismatches
			5;	Indels
				0;
				Gaps
				0;

```
Oy      1 TLWGEFOVAGRSARGGCOTSDAAAEALNASSKEEALOIRREKPEKYLFQFHNLNSL 60
        | : ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    111 TIENGEFQIGRSARGGQOTANDAAEALNASSKEEAMRIKEKPKELFQYHNLSNL 170
```

QY	61	DRIE	DKT	PEP	70
Db	171	DRIF	AKA	PEP	180

RESULT 5  
Q67574  
ID Q67574 PRELIMINARY; PRT; 361 AA

DT 01-NOV-1996 (TREMBlere). 01, Created)

DT 01-DEC-2001 (Tremblay, 19, Last annotation update)

GN ALL.

OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

RN [1] SEQUENCE FROM N 2

RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G., Maxwell D.D., Russell D.D.

## Cloning of the complete DNA genomes of four bean-infecting

RT particle acceleration.";  
 PI: phvtonathology 81:980-985(1991)

SEQUENCE FROM N. A.

RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Farla J.C.,  
RA Morales F.J., Maxwell D.P.,  
RT "Differentiation of bean-infecting geminiviruses by nucleic acid

RT hybridization probes and aspects of bean golden mosaic in Brazil.";  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.:  
 RT "Genetic diversity in geminiviruses causing bean golden mosaic  
 disease: The nucleotide sequence of the infectious cloned DNA  
 RT components of a Brazilian isolate of bean golden mosaic geminivirus.";  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M88686; AAA46312.1; -;  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 SO SEQUENCE 361 AA; 41041 MW; 009AC7ACAF06B788 CRC64;

Query Match 79.0%; Score 286; DB 12; Length 361;  
 Best Local Similarity 79.1%; Pred. No. 1.3e-24;  
 Matches 53; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

OY 4 MGEPVAGRSARGGCGTSDNDAAEALNASSKEEALQIRKIPKXYLFQFHNLSNLDRI 63  
 DB 113 WGHFOVDRSARGGQOTANDAAEALNASSKEEAMQIRKLPKFLFYHNLSSNLDRI 172  
 OY 64 FDKTPEP 70  
 DB 173 FTKAPDP 179

RESULT 6  
 O9QDB1  
 ID 09QDB1; PRELIMINARY; PRT; 225 AA.  
 AC 09QDB1;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID=69263;  
 RX [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CGMV-BR;  
 RA Faria J.C.;  
 RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from  
 RT Brazil.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF188708; AAF06318.1; -;  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 FT NON\_TER 225 225  
 SO SEQUENCE 225 AA; 25766 MW; 1089CBBD8D15B5D CRC64;

Query Match 78.5%; Score 284; DB 12; Length 225;  
 Best Local Similarity 79.1%; Pred. No. 1.2e-24;  
 Matches 53; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 4 MGEPVAGRSARGGCGTSDNDAAEALNASSKEEALQIRKIPKXYLFQFHNLSNLDRI 63  
 DB 113 WGHFOVDRSARGGQOTINDAASALNASSKEEAMQIRKLPKFLFYHNLSSNLDRI 172  
 OY 64 FDKTPEP 70  
 DB 173 FTKAPDP 179

RESULT 7  
 O9ELT8

ID O9ELT8; PRELIMINARY; PRT; 314 AA.  
 AC O9ELT8;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS Sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID=100755;  
 RX [1]  
 RN SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 RT States.";  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF288227; AAG01006.1; -;  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 SO SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 78.5%; Score 284; DB 12; Length 314;  
 Best Local Similarity 82.4%; Pred. No. 1.8e-24;  
 Matches 56; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 TLVWGEFVAGRSARGGCGTSDNDAAEALNASSKEEALQIRKIPKXYLFQFHNLSNLDRI 60  
 DB 110 TLVWGEFVAGRSARGGQOTANDAAEALNASSKEEALQIRKLPKXYLFQFHNLSNLDRI 169  
 OY 61 DRIFDKTP 68  
 DB 170 DRIFSPPE 177

RESULT 8  
 O9QSS5  
 ID 09QSS5; PRELIMINARY; PRT; 364 AA.  
 AC 09QSS5;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 GN AC1.  
 OS Sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID=100755;  
 RX [1]  
 RN SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 RT States.";  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl  
 RT virus : nucleotide sequence and phylogenetic relationships.";  
 RL Online Publication.  
 DR EMBL: AF104036; AAD47173.1; -;  
 DR InterPro: IPR001191; Gemin1\_AL1.  
 DR Pfam: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemin1\_AL1; 1.  
 SO SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;



```

ID 039180 PRELIMINARY; PRT; 234 AA.
AC 039180;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ACT (FRAGMENT).
OS Geminiviridae.
OC Viruses; ssDNA viruses.
OX NCBI_TaxID=10811;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOMATO STRAIN:
RA Guzman P., Arraondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
   Infecting Tomatoes in Venezuela."
RL Plant Dis. 81:312-312(1997).
DR EMBL: AF026553; AAB82605.1; -.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.
FT NON_TER 234
SQ SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;

```

```

Query Match 73.8%; Score 267; DB 12; Length 234;
Best Local Similarity 70.0%; Pred. No. 1,le-22;
Matches 49; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 1 TLWGCEFOVAGRSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLFOFHNLNSNL 60
   1 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 110 TIRMGTFQIDGRSARGGCGTANDAAEALNASSKEALQITREKIPKYLFOFHNLSSDL 169
OY 61 DRIFDKTPEP 70
   11111111
Db 170 DRIFMKAPNP 179

```

```

RESULT 13
O92089 PRELIMINARY; PRT; 190 AA.
AC 092089;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
   AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YOKOHAMA3;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001315; BAA34033.1; -.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.
FT NON_TER 190
SQ SEQUENCE 190 AA; 21432 MW; AAC093D1D1610PAD CRC64;

```

```

Query Match 72.9%; Score 264; DB 12; Length 190;
Best Local Similarity 61.2%; Pred. No. 2e-22;
Matches 52; Conservative 8; Mismatches 9; Indels 16; Gaps 1;
OY 1 TLWGCEFOVAGRSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLFOFHNLNSNL 60
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

```

```

Db 85 TLEMGTFQIDGRSARGGCGTANDACAEALNASSKABALITREKIPKDFIFQYHNLNSNL 144
OY 61 DRI-----FDKTP 69
   111-----11111
Db 145 DRIFAPLEVFVCPFTASSFDQVPE 169

```

```

RESULT 14
O92084 PRELIMINARY; PRT; 190 AA.
AC 092084;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
   AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YOKOHAMA5;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001318; BAA34039.1; -.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.
FT NON_TER 190
SQ SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

```

```

Query Match 72.9%; Score 264; DB 12; Length 190;
Best Local Similarity 61.2%; Pred. No. 2e-22;
Matches 52; Conservative 8; Mismatches 9; Indels 16; Gaps 1;

```

```

OY 1 TLWGCEFOVAGRSARGGCGTSDNDAAEALNASSKEALQITREKIPKYLFOFHNLNSNL 60
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 85 TLEMGTFQIDGRSARGGCGTANDACAEALNASSKABALITREKIPKDFIFQYHNLNSNL 144
OY 61 DRI-----FDKTP 69
   111-----11111
Db 145 DRIFAPLEVFVCPFTASSFDQVPE 169

```

```

RESULT 15
O9W827 PRELIMINARY; PRT; 190 AA.
AC 09W827;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORAL-1, PARTIAL AND
   COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GORAL;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001303; BAA34010.1; -.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.
DR PRINTS: PR00227; GEMCOATL1.
DR ProDom: PD000736; Gemini_AL1.

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:53 ; Search time 34.62 Seconds

(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346a-9

Perfect score: 362  
Sequence: 1 TLVWGEFQVAGRSARGCQT.....FOFHNLSNLDRIKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	356	98.3	352	1 VAL1_TGMV	P03567 tomato gold
2	256	70.7	361	1 VAL1_PYMV	P27258 potato yell
3	243	67.1	358	1 VAL1_CLVN	P14982 cassava lat
4	243	67.1	358	1 VAL1_CLVN	P14972 cassava lat
5	236	65.2	362	1 VAL1_TYLCU	P36279 tomato yell
6	229	63.3	359	1 VAL1_TYLCU	P38609 tomato yell
7	225	62.2	349	1 VAL1_PHVU	P06923 pepper huas
8	222	61.3	359	1 VAL1_TYLCU	P27260 tomato yell
9	220	60.8	353	1 VAL1_TYLCU	P05175 bean golden
10	219	60.5	355	1 VAL1_ABMV	P21947 abutilon mo
11	216	59.7	358	1 VAL1_BCTV	P14991 beet curly
12	216	59.7	361	1 VAL1_TMOV	P06657 tomato motc
13	207	57.2	357	1 VAL1_TYLCU	P27259 tomato yell
14	125	34.5	347	1 VAL1_SICV	P29048 squash leaf
15	64.5	17.8	299	1 Y175_HELPJ	O9ymq7 heliobacte
16	62.5	17.3	447	1 TBL2_HUMAN	O9y4p3 homo sapien
17	62	17.1	295	1 VAL1_TYDVA	P13617 tobacco yel
18	61	16.9	1483	1 CYP1_YEAS	P12351 saccharomyc
19	60.5	16.7	1713	1 LMA3_HUMAN	Q16787 homo sapien
20	60	16.6	319	1 LHA2_STAM	O99r35 staphylococ
21	60	16.6	447	1 CDSA_DROME	P56079 d phosphati
22	59	16.3	1287	1 SKI2_YEAS	P35207 saccharomyc
23	58.5	16.2	297	1 RRP2_RABVP	P06747 rabies viru
24	58.5	16.2	299	1 Y175_HELPJ	P56112 heliobacte
25	58.5	16.2	367	1 LHX4_MOUSE	P53776 mus musculu
26	58.5	16.2	589	1 YB67_HAETN	P45221 haemophilus
27	58.5	16.2	1044	1 BUB1_SCHPO	O94751 schizosacch
28	58	16.0	207	1 IL6_MARMO	O35736 marmota mon
29	58	16.0	256	1 YKJ9_YEAS	P34242 saccharomyc
30	58	16.0	617	1 YACH_ECOLI	P36682 escheichia
31	58	16.0	874	1 SUAP_BACLI	P49052 bacillus li
32	57.5	15.9	136	1 Y452_CAEEL	O62250 caenorhabdi
33	57	15.7	355	1 CRTB_RHOSH	P54905 rhodobacter

34	57	15.7	492	1 MOT3_MOUSE	O35308 mus musculu
35	57	15.7	492	1 MOT3_RAT	O70461 rattus norv
36	56.5	15.6	428	1 GBAL_CANAL	P28868 candida alb
37	56.5	15.6	470	1 RHBA_RHIME	O923r2 rhizobium m
38	56.5	15.6	511	1 HUTH_VIBCH	O9ksg4 vibrio chol
39	56	15.5	332	1 IPAD_SHIFL	P18013 shigella fl
40	56	15.5	336	1 Y625_METJA	Q58042 methanococc
41	56	15.5	432	1 PROA_DEIRA	O9rtd9 methanococc
42	56	15.5	513	1 HEMO_CHICK	P18080 gallus gall
43	56	15.5	601	1 CYSL_BUCAI	P57503 buchnera ap
44	56	15.5	795	1 LON_AQUAE	O66505 aquifex aeo
45	56	15.5	799	1 AFSK_STRCO	P54741 streptomyc

## ALIGNMENTS

RESULT 1	VAL1_TGMV	STANDARD:	PRT: 352 AA.
ID	VAL1_TGMV		
AC	P03567:		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	01-JUN-1994 (Rel. 29, Last annotation update)		
DE	ALI protein.		
GN	ACI.		
OS	Tomato golden mosaic virus (TGMV).		
OC	Viruses; ssDNA viruses; geminiviridae; Begomovirus.		
OX	NCBI_Taxid=10831;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Hamilton W.D.O., Steh V.E., Coutts R.H.A., Buck K.W.;		
RT	*Complete nucleotide sequence of the infectious cloned DNA components		
RT	of tomato golden mosaic virus: potential coding regions and regulatory		
RT	sequences.		
RL	EMBO J. 3:2197-2205 (1984).		
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
DR	EMBL; K02029; -; NOT_ANNOTATED_CDS.		
DR	PIR; A04170; QOCVLI.		
DR	InterPro; IPR001191; Gemin1_AL1.		
DR	Pfam; PF00799; Gemin1_AL1.1.		
DR	PRINTS; PR00227; GEMCOATNL.		
DR	ProDom; PD000736; Gemin1_AL1.1.		
FT	ProDom; PD000736; Gemin1_AL1.1.		
FT	NP_BIND 223 230 ATP (POTENTIAL).		
SQ	SEQUENCE 352 AA; 40332 MW; C33C938B964BAA4 CRC64;		
Query Match	98.3%; Score 356; DB 1; Length 352;		
Best Local Similarity	98.6%; Pred. No. 2.2e-32;		
Matches 69; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
QY	1 TLVWGEFQVAGRSARGCQTNDAAAFALNASSKEEALQIIRKIPKYLFOFHNLSNL 60		
DB	111 TLVWGEFQVAGRSARGCQTNDAAAFALNASSKEEALQIIRKIPKYLFOFHNLSNL 170		
QY	61 DRIDFKTPPEP 70		
DB	171 DRIDFKTPPEP 180		
RESULT 2	VAL1_PYMV	STANDARD:	PRT: 361 AA.
ID	VAL1_PYMV		

```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 protein.
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91311403; PubMed-1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RT potato yellow mosaic virus."
RL J. Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00940; BAA00782.1; -.
CC PIR; J00364; OOCVPT.
CC InterPro: IPR001191; Geminl_AL1.
CC Pfam; PF00799; Geminl_AL1; 1.
CC PRINTS; PR00227; GEMCOATALL.
CC ProDom; PD000736; Geminl_AL1; 1.
CC ATP-binding.
CC NP_BIND 222 229 ATP (POTENTIAL).
CC SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match 70.7%; Score 256; DB 1; Length 361;
Best Local Similarity 69.6%; Pred. No. 3.1e-21;
Matches 48; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFOVAGRSARGCGQTSNDAAAEALNASSKEALQIIRKIPKYLFPFHNLNSNL 60
DB 110 TLWGEFOVAGRSARGCGQTSNDAAAEALNASSKEALQIIRKIPKYLFPFHNLNSNL 169
QY 61 DRIFDKTPEP 69
DB 170 DRIFDKTPEP 178

RESULT 3
VAL1_CLVKN STANDARD; PRT; 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Gay M.R.;
RT "Nucleotide sequence of cassava latent virus DNA."
RL Nature 301:260-262(1983).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J02057; -. NOT_ANNOTATED_CDS.
CC InterPro: IPR001191; Geminl_AL1.
CC Pfam; PF00799; Geminl_AL1; 1.
CC PRINTS; PR00227; GEMCOATALL.
CC ProDom; PD000736; Geminl_AL1; 1.
CC ATP-binding.
CC NP_BIND 220 227 ATP (POTENTIAL).
CC SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 67.1%; Score 243; DB 1; Length 358;
Best Local Similarity 62.9%; Pred. No. 8.6e-20;
Matches 44; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 1 TLWGEFOVAGRSARGCGQTSNDAAAEALNASSKEALQIIRKIPKYLFPFHNLNSNL 60
DB 109 TLWGEFOVAGRSARGCGQTSNDAAAEALNASSKEALQIIRKIPKYLFPFHNLNSNL 168
QY 61 DRIFDKTPEP 70
DB 169 DRIFDKTPEP 178

RESULT 4
VAL1_CLVKN STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90174930; PubMed-2308831;
RA Morris B., Coates J., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain)."
RL Nucleic Acids Res. 18:197-198(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X17095; CAA34953.1; -.
CC PIR; S07594; S07594.
CC InterPro: IPR001191; Geminl_AL1.
CC Pfam; PF00799; Geminl_AL1; 1.
CC PRINTS; PR00227; GEMCOATALL.
CC ProDom; PD000736; Geminl_AL1; 1.
CC ATP-binding.
CC NP_BIND 220 227 ATP (POTENTIAL).
CC SEQUENCE 358 AA; 40435 MW; 1DB16BB0CB2D5E2C CRC64;

Query Match 67.1%; Score 243; DB 1; Length 358;
Best Local Similarity 62.9%; Pred. No. 8.6e-20;
Matches 44; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 1 TLWGEFOVAGRSARGCGQTSNDAAAEALNASSKEALQIIRKIPKYLFPFHNLNSNL 60
DB 109 TLWGEFOVAGRSARGCGQTSNDAAAEALNASSKEALQIIRKIPKYLFPFHNLNSNL 168

```

QY 61 DRIFDKTPEP 70  
 DB 169 DRIFGEPPAP 178

## RESULT 5

VAL1\_TYLCU STANDARD: PRT: 362 AA.  
 AC P36279;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36447;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93139778; PubMed=8423446;  
 RA Dry I.B., Ridgen J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;  
 RT "Nucleotide sequence and genome organization of tomato leaf curl  
 geminivirus." 74:147-151(1993).  
 RL J. Gen. Virol. 74:147-151(1993).  
 CC -i SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 DR PIR: J01887; J01887.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.  
 KW ATP-binding.  
 FT NP\_BIND 221  
 SQ SEQUENCE 362 AA; 41197 MW; 343E7184B4704098 CRC64;

Query Match 65.2%; Score 236; DB 1; Length 362;  
 Best Local Similarity 55.3%; Pred. No. 5.2e-19;  
 Matches 47; Conservative 9; Mismatches 13; Indels 16; Gaps 1;

QY 1 TLWGEFOVAGSARGCGQTSNDAAAEALNMSKEALQITREKIPKYLFOFHNLNSND 60  
 DB 110 TLWGEFOIDGRSARGGQSDANDAYQAALNMGSKSEALNVLREAPKDYVLFPHNLNSL 169  
 QY 61 DRI-----FDKTPPE 69  
 DB 170 DRIETPPELVYVSPFLSSFDKPE 194

## RESULT 6

VAL1\_TYLCU STANDARD: PRT: 359 AA.  
 AC P38609;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE AL1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=37139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94256836; PubMed=8198442;  
 RA Norris E., Hidalgo E., Accotto G., Moriones E.;  
 RT "High similarity among the tomato yellow leaf curl virus isolates  
 from the west Mediterranean basin: the nucleotide sequence of an  
 infectious clone from Spain." Arch. Virol. 135:165-170(1994).  
 RL -i SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z25751; CAA81026.1; -  
 DR PIR: S39211; S39211.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.  
 KW ATP-binding.  
 FT NP\_BIND 221  
 SQ SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 63.3%; Score 229; DB 1; Length 359;  
 Best Local Similarity 60.9%; Pred. No. 3.1e-18;  
 Matches 42; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 2 LWGEFOVAGSARGCGQTSNDAAAEALNMSKEALQITREKIPKYLFOFHNLNSND 61  
 DB 111 LWGEFOIDGRSARGGQSDANDAYAKAINAGSKSEALDVYIKELAPRDYLIHFNHNSND 170  
 QY 62 RIFDKTPEP 70  
 DB 171 RVFOVPPAP 179

## RESULT 7

VAL1\_PHVU STANDARD: PRT: 349 AA.  
 AC Q06923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein.  
 GN AL1.  
 OS Pepper huasteco virus (PHV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=28349;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94015007; PubMed=8409944;  
 RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,  
 KA Rivera-Bustamante R.F.;  
 RT "Complete nucleotide sequence of pepper huasteco virus: analysis and  
 comparison with bipartite geminiviruses." J. Gen. Virol. 74:2225-2231(1993).  
 RL -i SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X70418; CAA49856.1; -  
 DR PIR: S31875; S31875.  
 DR PIR: J02300; J02300.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.  
 KW ATP-binding.  
 FT NP\_BIND 221  
 SQ SEQUENCE 349 AA; 39722 MW; D5F4E76CD56370F4 CRC64;

Query Match 62.2%; Score 225; DB 1; Length 349;

Best Local Similarity 60.0%; Pred. No. 8,4e-18;  
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLWGEFQVAGRSARGCGGCTNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSND 60  
110 TLWGEFQVAGRSARGCGGCTNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSND 169

QY 61 DRIFDKTPEP 70  
170 NRIFQTPPEP 179

Db

RESULT 8  
ID VAL1\_TYLCM STANDARD; PRT; 359 AA.  
AC P27260;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 protein (C1 protein).  
GN C1.  
OS Tomato yellow leaf curl virus (strain Marmende) (TYLCV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-9210760; PubMed-1840676;  
RA Kheyri-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,  
RT Groenbom B.;  
RT "Tomato yellow leaf curl virus from Sardinia is a  
RT whitefly-transmitted monopartite geminivirus";  
RT Nucleic Acids Res. 19:6763-6769(1991).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X61153; CAA3466.1; -  
CC PIR: S22593; S22593.  
CC InterPro: IPR001191; Gemin1\_AL1.  
CC Pfam: PF00799; Gemin1\_AL1; 1.  
CC PRINTS: PR00227; GEMCOATALL.  
CC ProDom: PD000736; Gemin1\_AL1; 1.  
CC ATP-binding. 220 ATP (POTENTIAL).  
FT NP-BIND 220  
SQ SEQUENCE 359 AA; 40733 MW; 9717BA407C93EFA7 CRC64;

Query Match 61.3%; Score 222; DB 1; Length 359;  
Best Local Similarity 58.0%; Pred. No. 1.9e-17;  
Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 2 LWGEFQVAGRSARGCGGCTNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSND 61  
111 LWGEFQVAGRSARGCGGCTNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSND 170

QY 62 RIFDKTPEP 70  
171 KVFQVPEP 179

Db

RESULT 9  
ID VAL1\_BGMV STANDARD; PRT; 353 AA.  
AC P05175;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE AL1 protein (40.2 kDa protein).  
GN Acl.  
OS bean golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;  
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene  
RT regulation in geminiviruses";  
RT Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M10070; AAA6318.1; -  
CC InterPro: IPR001191; Gemin1\_AL1.  
CC Pfam: PF00799; Gemin1\_AL1; 1.  
CC PRINTS: PR00227; GEMCOATALL.  
CC ProDom: PD000736; Gemin1\_AL1; 1.  
CC ATP-binding. 222 ATP (POTENTIAL).  
FT NP-BIND 222  
SQ SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match 60.8%; Score 220; DB 1; Length 353;  
Best Local Similarity 60.0%; Pred. No. 3.1e-17;  
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLWGEFQVAGRSARGCGGCTNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSND 60  
110 TLWGEFQVAGRSARGCGGCTNDAAEALNASSKEEALQIRKIPKYLFOFHNLNSND 169

QY 61 DRIFDKTPEP 70  
170 ERIFKVPPEP 179

Db

RESULT 10  
ID VAL1\_ABMV STANDARD; PRT; 355 AA.  
AC P21947;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 protein.  
GN Acl.  
OS Abutilon mosaic virus (isolate West India).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91020984; PubMed-2219703;  
RA Frischmuth T., Zimmet G., Jeske H.;  
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
RT as well as eukaryotic features";  
RT Virology 178:461-468(1990).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: X15983; -: NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; QOCVW1.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 DR ATP-binding.  
 KW NP\_BIND 221 ATP (POTENTIAL).  
 FT SEQUENCE 355 AA; 40257 MW; 16A2CA8A63251E95 CRC64;  
 SQ

Query Match 60.5%; Score 219; DB 1; Length 355;  
 Best Local Similarity 58.6%; Pred. No. 4e-17;  
 Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLWGEFQVAGRSARGGCGTSDNDAAEALNASSKEEALQIIRKIPKTYLFOFHNLSNL 60  
 DB 110 TLWGEFQIDGRSARGGCGTANDSYAKALNAGDVOSALNIRKEQPKDYVLQHNIRSNL 169  
 OY 61 DRIPDKTEPP 70  
 DB 170 ERIFAKAPEP 179

RESULT 11  
 VAL1\_BCTV STANDARD; PRT; 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE A1 protein (40.8 kDa protein).  
 OS Beet curly top virus (BCTV).  
 CC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
 OX NCBI\_TaxID=10840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 RT beet curly top virus."  
 RL EMBL J. 5:1761-1767(1986).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04144; -: NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND 222 ATP (POTENTIAL).  
 FT SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;  
 SQ

Query Match 59.7%; Score 216; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 8.7e-17;  
 Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLWGEFQVAGRSARGGCGTSDNDAAEALNASSKEEALQIIRKIPKTYLFOFHNLSNL 60  
 DB 110 TLWGEFQIDGRSARGGCGTANDSYAKALNATSLDQALQILKEQPKDYVLQHNLSNL 169  
 OY 61 DRIPDKTEPP 70  
 DB 170 OKIFORPPDP 179

RESULT 12  
 VAL1\_TMOV STANDARD; PRT; 361 AA.  
 ID VAL1\_TMOV  
 AC Q06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein.  
 GN A11.  
 OS Tomato mottle virus (Isolate Florida) (TMOV).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93107858; PubMed=1469361;  
 RA Abouzaid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 RT isolated from tomatoes in Florida."  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES A11 PROTEIN FAMILY.  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: L14460; AAC32414.1; -;  
 DR PIR: J01870; J01870.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND 222 ATP (BY SIMILARITY).  
 FT SEQUENCE 361 AA; 40516 MW; 8138B65CEBAC6550 CRC64;  
 SQ

Query Match 59.7%; Score 216; DB 1; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 8.8e-17;  
 Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLWGEFQVAGRSARGGCGTSDNDAAEALNASSKEEALQIIRKIPKTYLFOFHNLSNL 60  
 DB 110 TLWGEFQIDGRSARGGCGTANDSYAKALNAGSVOSALNIRKEQPKDYVLQHNIRSNL 169  
 OY 61 DRIPDKTEPP 70  
 DB 170 ERIFAKAPEP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT; 357 AA.  
 ID VAL1\_TYLCV  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RT with a single genomic component.";  
 RL Virology 185:151-161(1991).

```

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X15656; CAA33688.1; -.
CC PIR: D40779; Q0CVCL.
CC InterPro: IPR001191; Gemini1_AL1.
CC Pfam: PF00799; Gemini_AL1; 1.
CC PRINTS: PR00227; GEMCOATAL1.
CC ProDom: PD000736; Gemini_AL1; 1.
CC ATP-binding. 219
CC NP_BIND
CC SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;

Query Match
Best Local Similarity 57.2%; Score 207; DB 1; Length 357;
Matches 40; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 WGEFVAGRSARGCGCTSDAAAEALNASSKEEALQIREKIPKYLFPFHNLNSNDR 63
DB 111 FGYSQIDGHSARGCGOSANDAYAEALNASSKEEALNILEKPKYIILQFHLNSNDR 170

QY 64 F 64
DB 171 F 171

RESULT 14
VAL1_SLCV STANDARD; PRT; 347 AA.
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE AL1 protein.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-91082449; PubMed-1984668;
RA Lazarowitz S.G.; Lazdins I.B.;
*Infectivity and complete nucleotide sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype."
RT Virology 180:58-69(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M38183; AAC32410.1; AL1_INTR.
CC PIR: C36785; Q0CVS1.
CC InterPro: IPR001191; Gemini1_AL1.
CC Pfam: PF00799; Gemini_AL1; 1.
CC PRINTS: PR00227; GEMCOATAL1.
CC ProDom: PD000736; Gemini_AL1; 1.
CC ATP-binding. 218
CC NP_BIND
CC SEQUENCE 347 AA; 39110 MW; AFDAEBDD122110E CRC64;

```

```

Query Match
Best Local Similarity 34.5%; Score 125; DB 1; Length 347;
Matches 25; Conservative 15; Mismatches 22; Indels 4; Gaps 1;

QY 5 GEFOVAGRSARGCGCTSDAAAEALNASSKEEALQIREKIPKYLFPFHNLNSNDR 64
DB 116 GRYVRS-----GSKNKDDVYHNNAVNSAGAEALDIKADPKRTFIVNHLNVERLRF 171

QY 65 DKTEPP 70
DB 172 QKPEP 177

RESULT 15
Y175_HELPJ STANDARD; PRT; 299 AA.
ID Y175_HELPJ
AC Q9ZMO7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein JHP0161 precursor.
GN JHP0161.
OS Helicobacter pylori 399 (Campylobacter pylori 399).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99120557; PubMed-9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voris G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- SIMILARITY: BELONGS TO THE PPTC/PARVULIN FAMILY OF ROTAMASES.
CC -----
CC STRONG, TO C. JEJUNI CHF2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001454; AAD05744.1; -.
CC HSSP: Q13526; IPIN.
CC DR InterPro: IPR000297; Rotamase.
CC DR Pfam: PF00639; Rotamase; 1.
CC DR PROSITE: PS01096; PPTC_PPTASE_1; 1.
CC DR PROSITE: PS50198; PPTC_PPTASE_2; 1.
CC KW Hypothetical protein; Isomerase; Rotamase; signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 299 POTENTIAL.
FT DOMAIN 154 253 HYPOTHETICAL PROTEIN JHP0161.
FT SEQUENCE 299 AA; 34040 MW; 9C037B1CD110143 CRC64;

Query Match
Best Local Similarity 17.8%; Score 64.5; DB 1; Length 299;
Matches 18; Conservative 14; Mismatches 21; Indels 2; Gaps 2;

QY 22 NDAAEALNLS-----SKEEALQIREKIPKYLFPFHNLNSNDR 62
DB 93 NEAKAEKLNQTPERKAMMEAVKQALVEFWAKQAEVKKIQIPKEMQFV--NANKQD 150

QY 63 IFDK 66
DB 151 LEVR 154

```

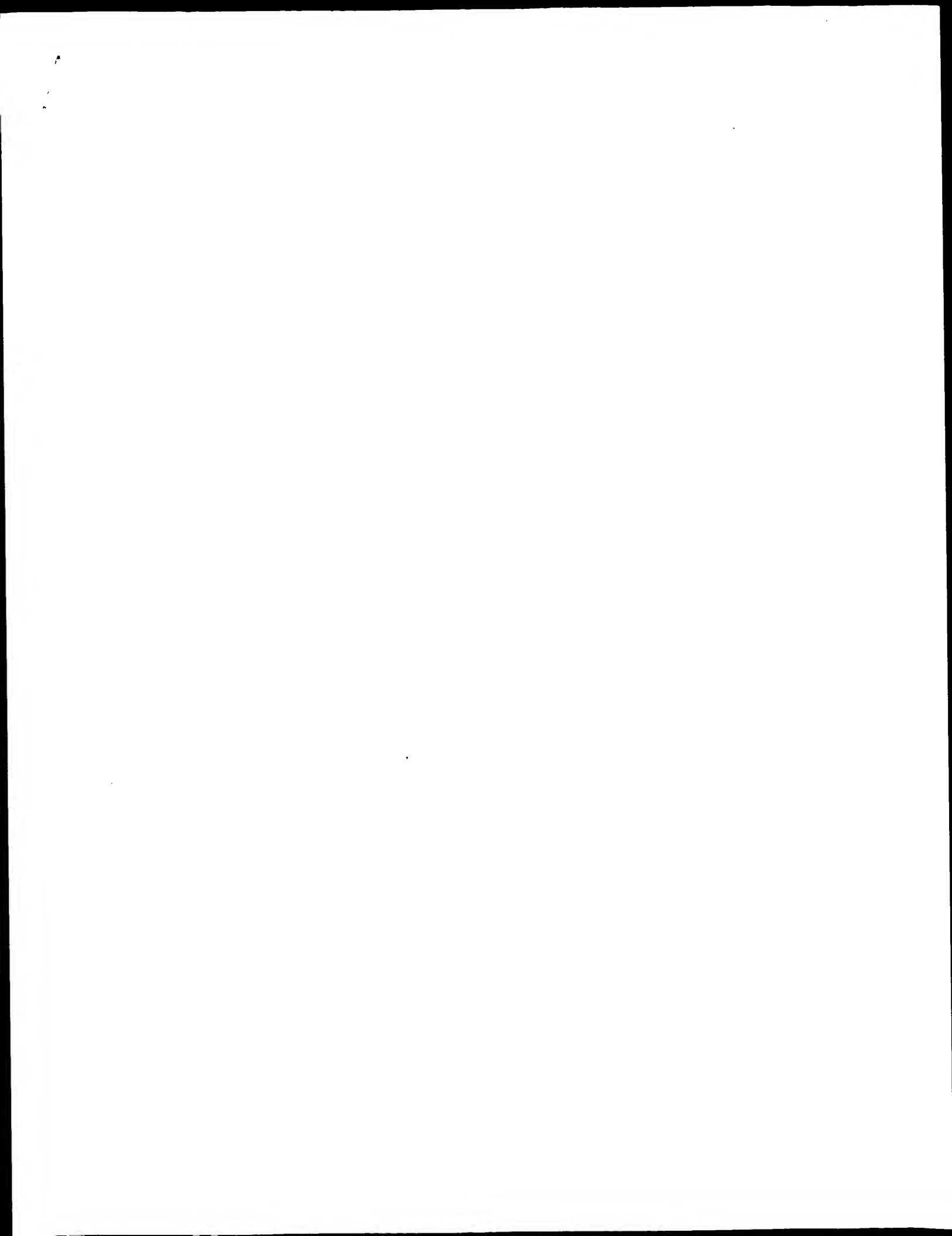
Thu Oct 10 16:06:19 2002

us-09-289-346a-9.rsp

Page 7

Search completed: October 10, 2002, 02:09:53  
Job time: 1351 sec

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:45 ; Search time 69.63 Seconds  
(without alignments) 96.600 Million cell updates/sec

Title: US-09-289-346A-9

Perfect score: 362

Sequence: 1 TLVWGEFQVAGRSARGCQOT.....FFPHNLSNLDRIFFDKTPEP 70

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	356	98.3	352	1	QOCVCL1
2	256	70.7	361	1	QOCVPT
3	243	67.1	358	2	S07594
4	236	65.2	362	1	J01887
5	229	63.3	359	2	S39211
6	225	62.2	349	2	J02300
7	225	62.2	349	2	S31875
8	222	61.3	359	2	S22593
9	219	60.5	351	2	J02327
10	219	60.5	351	2	QOCVW1
11	216	59.7	358	1	J01870
12	216	59.7	359	2	S39235
13	216	59.7	385	2	S28360
14	209	57.7	360	2	S59885
15	207	57.2	357	1	QOCVCL1
16	142	39.2	131	2	S45059
17	125	34.5	347	1	QOCVSL1
18	68	18.8	587	1	J01419
19	67	18.5	840	2	T36175
20	64.5	17.8	299	2	B71967
21	63.5	17.5	642	2	D90558
22	63	17.4	714	2	C95382
23	62.5	17.3	447	2	T12544
24	62	17.1	154	2	AD3475
25	62	17.1	181	2	G97976
26	62	17.1	295	2	D42452
27	62	17.1	340	2	B97011
28	61.5	17.0	201	2	A81380
29	61.5	17.0	481	2	A70091

## ALIGNMENTS

## RESULT 1

QOCVCL1

AL1 protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C>Date: 28-Aug-1985 #sequence, revision 28-Aug-1985 #text, change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of tomato

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 98.3%; Score 356; DB 1; Length 352;

Best Local Similarity 98.6%; Pred. No. 2.5e-32;

Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TLVWGEFQVAGRSARGCQOTSDAAAEALNASSKEEALQIRREKIPKYLFFPHNLSNL 60

DB 111 TLVWGEFQVAGRSARGCQOTSDAAAEALNASSKEEALQIRREKIPKYLFFPHNLSNL 170

QY 61 DRIFDKTPEP 70

DB 171 DRIFDKTPEP 180

## RESULT 2

QOCVPT

AL1 protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C>Date: 30-Jun-1992 #sequence, revision 30-Jun-1992 #text, change 16-Jun-2000

C:Accession: J00364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: J00362; MIMD:91311403

A:Accession: J00364

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-361 <COU>

A:Cross-references: GB:D00940; NID:9222458; PIDN:BA00782.1; PID:9222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein





```

QY      1  TLWNGCEPQVARSNRAGGCGOTSDNDAAEALNMSKKEALQIIRREKIPERKYLFOFHNLSNL 60
        | : ||| ||| ||| ||| ||| ||| | : | : ||| ||| ||| ||| ||| ||| ||| |||
Db      107  TIEMGDFIDGSRAGGQSDANDSTAKALNAGSVOSALAVLREPOPKDFVLONHNIRSNL 166

QY      61  DRIFDKTREP 70
        : ||| | |||
Db      167  ERIFAKAPEP 176

RESULT  12
S39235
gene C1 protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C:Accession: S39235
R:Crespi, S.; Norris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity
A:Reference number: S39233
A:Accession: S39235
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <CRD>
A:Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C:Superfamily: tomato golden mosaic virus AL1 protein

```

Query Match	59.78;	Score 216;	DB 2;	Length 359;
Best Local Similarity	56.58;	Pred. No. 1.4e-16;		
Matches	39;	Conservative 12;	Mismatches 18;	Indels 0;
				Gaps 0

[illegible]

```

RESULT      13
S28360
ALI protein - beet curly top virus
C:Species: beet curly top virus
C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C:Accession: S28360
R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus
A:Reference number: S28360
A:Accession: S28360
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-385 <STRA>
A:Cross-references: GB:J24597; EMBL:X04144; NID:g210678; PIDN:AAAA2751.1; PID:g210679
;Superfamily: tomato golden mosaic virus AL1 protein

```

Query Match	59.7%;	Score 216;	DB 2;	Length 385;
Best Local Similarity	55.7%;	Pred. NO. 1.5e-16;		
Matches	39;	Conservative	16;	Mismatches 15;
				Indels 0;
				Gaps 0;

```
QY      1 TLWGEFOVAGRSRGGCOTSDNAAABAALNNSSKEALDIIIREK IEPKYIFOFNNLSNL 60
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      137 TTEMEFEIDGSGAGGQIANDSTAKALNNTSLDQALDITKEBQRPDYDGLDHHNLNNA 156
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61 DRIEDKTPRP 70
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      197 QKTFQRPDP 206
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

♦ RESULT 14

S59885  
replication-associated protein C1 - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
C:Accession: S59885  
R:Hong, Y.; Harrison, B. D.  
Submitted to the EMBL Data Library, February 1995  
A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
d geminiviruses.  
A:Reference number: S58346  
A:Accession: S59885  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <HON>  
A:Cross-References: EMBL:Z48182, NID:G944838, PIDN:CMA88229.1, PID:9974211  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	57.7%;	Score 209;	DB 2;	Length 360;
Best Local Similarity	60.6%;	Pred. No. 8.7e-16;		
Matches 40;	Conservative 12;	Mismatches 14;	Indels 0;	Gaps 0;

QY 4 WGEFVAGRGARGGCOTSDNAAAEALNASSKEEALQITREKIPERYLLPFQHNLNLSNDRI 63  
: || : ||||| : ||| : | | : ||| : : ||||| : |||  
Db 113 FGVRQIDGRARGGGOSANDAYAEALINSGSKAALDIIEKAPKQFVLQFNHLNANNDRI 172

QY	64	FDKTP	69
		:	
Db	173	FTPSAE	178

RESULT 15

AL1 protein - tomato yellow leaf curl virus  
 N:Alternate names: CI protein  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin-  
 A:Reference number: A40779; M01D:92024070  
 A:Accession: D40779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 A:Cross-References: GB:X15656; NID:g62204; PIDN:CAA33668.1; PID:g62207  
 A:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	57.2%;	Score 207;	DB 1;	Length 357;
Best Local Similarity	65.68;	Pred. No. 1.5e-15;		
Matches 40; Conservative	10;	Mismatches 11;	Indels 0;	Gaps 0.

```

0Y      4 WGEFVAGRSARGGCOTSNDAAEALNASSKEELQIREKIPREYLLFOFHNLNSNDRI 63
      : | : | | | | | : | | | | | : | | | | | : | : | | | | | | |
Db     111 FGVSGIDGRSARGGQGSANDAYAEALNSGSKSEALNIIEKAPRYDYLQFHNLSSNDRI 170

```

QY	64	F	64
		-	
Db	171	F	171

Search completed: October 10, 2002, 01:48:45  
Job time: 563 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:19 ; Search time 133.15 Seconds  
(without alignments)  
58.394 Million cell updates/sec

Title: US-09-289-346a-9  
Perfect score: 362  
Sequence: 1 TLWGGEFYVAGRSARGGCOT.....FGFHLNSNIDRIFDKTPRP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A\_Geneseq\_032802.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	362	100.0	70	21	AA18685	Mutant peptide der
2	356	98.3	70	21	AA18687	Peptide fragment f
3	356	98.3	356	21	AA18687	Amino acid sequenc
4	345	95.3	70	21	AA18688	Mutant peptide der
5	344	95.0	70	21	AA18692	Mutant peptide der
6	342	94.5	70	21	AA18684	Mutant peptide der
7	342	94.5	70	21	AA18690	Mutant peptide der
8	342	94.2	70	21	AA18678	Mutant peptide der
9	341	94.2	70	21	AA18686	Mutant peptide der
10	340	93.9	70	21	AA18689	Mutant peptide der
11	338	93.4	70	21	AA18680	Mutant peptide der

Result No.	Score	Query Match	Length	DB ID	Description	
12	338	93.4	70	21	AA18691	Mutant peptide der
13	335	92.5	70	21	AA18681	Mutant peptide der
14	334	92.3	70	21	AA18683	Mutant peptide der
15	332	91.7	70	21	AA18682	Mutant peptide der
16	324	89.5	70	21	AA18679	Mutant peptide der
17	322	61.3	353	18	AAW34338	Bean golden mosaic
18	322	61.3	353	18	AAW34332	Bean golden mosaic
19	322	61.3	353	18	AAW34333	Bean golden mosaic
20	322	61.3	353	18	AAW34334	Bean golden mosaic
21	322	61.3	353	18	AAW34335	Bean golden mosaic
22	322	61.3	359	17	AAW34325	Sardinian tomato y
23	322	61.3	359	17	AAW34326	Sardinian tomato y
24	322	61.3	359	17	AAW34327	Sardinian tomato y
25	320	60.8	353	8	AAW34328	ORF 4 gene product
26	320	60.8	361	18	AAW34329	Tomato mottle viru
27	320	60.8	361	18	AAW34330	Tomato mottle viru
28	320	60.8	361	18	AAW34331	Tomato mottle viru
29	320	60.8	361	18	AAW34332	Tomato mottle viru
30	313.5	59.0	361	8	AAW34333	Product of ORF 4 f
31	308	57.5	362	19	AAW56495	Tobacco leaf curl
32	307	57.2	357	18	AAW34329	Tomato yellow leaf
33	307	57.2	357	18	AAW34330	Tomato yellow leaf
34	307	57.2	357	18	AAW34331	Tomato yellow leaf
35	307	57.2	357	18	AAW34332	Tomato yellow leaf
36	307	57.2	357	18	AAW34333	Tomato yellow leaf
37	307	57.2	357	18	AAW34334	Tomato yellow leaf
38	307	57.2	357	18	AAW34335	Tomato yellow leaf
39	307	57.2	357	18	AAW34336	Tomato yellow leaf
40	307	57.2	357	18	AAW34337	Tomato yellow leaf
41	307	57.2	357	18	AAW34338	Tomato yellow leaf
42	307	57.2	357	18	AAW34339	Tomato yellow leaf
43	307	57.2	357	18	AAW34340	Tomato yellow leaf
44	307	57.2	357	18	AAW34341	Tomato yellow leaf
45	307	57.2	357	18	AAW34342	Tomato yellow leaf

## ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description	
1	362	100.0	70	21	AA18685	Mutant peptide der
2	356	98.3	70	21	AA18687	Peptide fragment f
3	356	98.3	356	21	AA18687	Amino acid sequenc
4	345	95.3	70	21	AA18688	Mutant peptide der
5	344	95.0	70	21	AA18692	Mutant peptide der
6	342	94.5	70	21	AA18684	Mutant peptide der
7	342	94.5	70	21	AA18690	Mutant peptide der
8	342	94.2	70	21	AA18678	Mutant peptide der
9	341	94.2	70	21	AA18686	Mutant peptide der
10	340	93.9	70	21	AA18689	Mutant peptide der
11	338	93.4	70	21	AA18680	Mutant peptide der

```

XX PS Transgenic plants with increased resistance to geminivirus infection
PT PT Comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX PS
XX PS Claim 53; Page 46; 73pp; English.
XX CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX S0 Sequence 70 AA;

Query Match          100.0%; Score 362; DB 21; Length 70;
Best local Similarity 100.0%; Pred. No. 8.2e-38;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLVWGEFQVAGRSARGCQTSNDAAAEALNASSKEEALQITREKIPKYLFGPHNLNSNL 60
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 tlwgefyvdgrsarqgcqtsndaaaealnasskeaalqitrekipekylfgfhlnlnsl 60

QY 61 DRIFDKTPEP 70
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 drifdktp 70

RESULT 2
AAB18677 standard; peptide; 70 AA.
XX ID AAB18677 standard; peptide; 70 AA.
XX AC AAB18677;
XX DT 22-JAN-2001 (first entry)
XX DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX OS Tomato golden mosaic virus.
XX PN WO200054573-A1.
XX PD 21-SEP-2000.
XX PF 15-MAR-2000; 2000WO-US06759.
XX PR 18-MAR-1999; 99US-0125004.
XX PR 09-APR-1999; 99US-0289346.
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX DR WPI; 2000-618851/59.
XX PT Transgenic plants with increased resistance to geminivirus infection
PT PT Comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT

```

```

XX PS Disclosure; Page 18; 73pp; English.
XX CC The present sequence is derived from a geminivirus replication (Rep)
XX protein, which is also known as AL1. AL1 binds double-stranded DNA,
XX catalyzes cleavage and ligation of single-stranded DNA, and interacts
XX with other viral and host proteins. Mutants of the AL1 protein are used
XX to produce transgenic plants. The mutation in AL1 is present in a
XX ribosome binding region, and expression of mutant AL1 protein imparts
XX increased resistance to geminivirus infection in the plant. Mutant AL1
XX proteins are useful for producing plants having increased resistance or
XX reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
XX tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
XX virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
XX yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
XX squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
XX beet curly top virus.
XX S0 Sequence 70 AA;

Query Match          98.3%; Score 356; DB 21; Length 70;
Best local Similarity 98.6%; Pred. No. 4.6e-37;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TLVWGEFQVAGRSARGCQTSNDAAAEALNASSKEEALQITREKIPKYLFGPHNLNSNL 60
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 tlwgefyvdgrsarqgcqtsndaaaealnasskeaalqitrekipekylfgfhlnlnsl 60

QY 61 DRIFDKTPEP 70
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 drifdktp 70

RESULT 3
AAB18687 standard; peptide; 356 AA.
XX ID AAB18687 standard; peptide; 356 AA.
XX AC AAB18687;
XX DT 22-JAN-2001 (first entry)
XX DE Amino acid sequence of a geminivirus replication protein of TGMV.
XX KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX OS Tomato golden mosaic virus.
XX PN WO200054573-A1.
XX PD 21-SEP-2000.
XX PF 15-MAR-2000; 2000WO-US06759.
XX PR 18-MAR-1999; 99US-0125004.
XX PR 09-APR-1999; 99US-0289346.
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX DR WPI; 2000-618851/59.
XX PT Transgenic plants with increased resistance to geminivirus infection
PT PT Comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT

```

PS Disclosure: Page 47-48; 73pp; English.  
 CC The present sequence represents a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 CC  
 CC Sequence 356 AA;  
 SO

Query Match 98.3%; Score 356; DB 21; Length 356;  
 Best Local Similarity 98.6%; Pred. No. 3.6e-36;  
 Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TLVWGEFVAGRSARGCGQTSNDAAAEALNASSKEEALQIIRKIPKYLPOFHNLNSL 60  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 1 ltvwgefvdgrrsargcgqtsndaaaealnasskeeaqlirkipkylfqfhnlnsl 169  
 OY 61 DRIFDKTPEP 70  
 |||||||  
 Db 170 drifdktp 179

RESULT 4  
 AAB18688  
 ID AAB18688 standard; peptide: 70 AA.  
 AC AAB18688;  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 19  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 20  
 FT /note= "wild type residue replaced with Ala"  
 FT  
 FT WO200054573-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000MO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region

PT  
 XX Disclosure: Page 48; 73pp; English.  
 PS  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 CC Sequence 70 AA;  
 SO

Query Match 95.3%; Score 345; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 1.1e-35;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TLVWGEFVAGRSARGCGQTSNDAAAEALNASSKEEALQIIRKIPKYLPOFHNLNSL 60  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 1 ltvwgefvdgrrsargcgqtsndaaaealnasskeeaqlirkipkylfqfhnlnsl 60  
 OY 61 DRIFDKTPEP 70  
 |||||||  
 Db 61 drifdktp 70

RESULT 5  
 AAB18692  
 ID AAB18692 standard; peptide: 70 AA.  
 AC AAB18692;  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 66  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 69  
 FT /note= "wild type residue replaced with Ala"  
 FT  
 FT WO200054573-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000MO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 DR WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure: Page 50; 73pp; English.

CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:

Query Match 95.0%; Score 344; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 1.5e-35;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVGEFQVAGRSARGCQTSNDAAAEALNASSKEBALQIREKIPKYLFGFHNLSNL 60  
 Db 1 tlwgeafqvdgrsargcqtndaaaealnasskeaalqirekipekylfghnlnsl 60  
 QY 61 DRFDKTPPEP 70  
 Db 61 drfdktpep 70

## RESULT 6

AAB18684  
 ID AAB18684 standard; peptide; 70 AA.

AC AAB18684;  
 XX  
 DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX

OS Synthetic.  
 OS Tomato golden mosaic virus.

XX  
 FH Key Location/Qualifiers

FT MISC-difference 7 /note= "wild type residue replaced with Ala"

FT MISC-difference 8 /note= "wild type residue replaced with Ala"

FT MISC-difference 30 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000MO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 DR

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 52; Page 45; 73pp; English.

CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA:

Query Match 94.5%; Score 342; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 2.6e-35;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVGEFQVAGRSARGCQTSNDAAAEALNASSKEBALQIREKIPKYLFGFHNLSNL 60  
 Db 1 tlwgeaavdgrsargcqtndaaaealnasskeaalqirekipekylfghnlnsl 60  
 QY 61 DRFDKTPPEP 70  
 Db 61 drfdktpep 70

## RESULT 7

AAB18690  
 ID AAB18690 standard; peptide; 70 AA.

AC AAB18690;  
 XX  
 DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX

OS Synthetic.  
 OS Tomato golden mosaic virus.

XX  
 FH Key Location/Qualifiers

FT MISC-difference 27 /note= "wild type residue replaced with Ala"

FT MISC-difference 30 /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000MO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 DR



```

XX  WPI: 2000-618651/59.
DR  Transgenic plants with increased resistance to geminivirus infection
XX  comprise a nucleic acid construct containing a nucleic acid sequence
PT  encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX  Disclosure: Page 49; 73pp; English.
XX
XX  The present sequence represents a mutant peptide, derived from a
XX  geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX  double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX  DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX  protein are used to produce transgenic plants. The mutation in AL1 is
XX  present in a ribosome binding region, and expression of mutant AL1
XX  protein imparts increased resistance to geminivirus infection in the
XX  plant. Mutant AL1 proteins are useful for producing plants having
XX  increased resistance or reduced sensitivity to a geminivirus such as
XX  tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX  virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX  cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX  virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX  virus, cotton leaf curl virus or beet curly top virus.
XX
XX  Sequence 70 AA:
XX
XX  Query Match 94.5%; Score 342; DB 21; Length 70;
XX  Best Local Similarity 95.7%; Pred. No. 2.6e-35;
XX  Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVAGRSARGCOTSDNDAAEALNASSKEEALQIRREKIPKYLQFHNLSNL 60
DB 1 tlwgefqvdsrargscqtsndaaalaaasskeaalqirrekipkyllqfhnlnsl 60
QY 61 DRIEDKTPPEP 70
DB 61 drifdktppep 70

RESULT 8
AAB18678
ID AAB18678 standard; peptide: 70 AA.
XX
XX AAB18678:
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX
XX MISC-difference 12 /note= "wild type residue replaced with Ala"
XX
XX MISC-difference 13 /note= "wild type residue replaced with Ala"
XX
XX MISC-difference 15 /note= "wild type residue replaced with Ala"
XX
XX W0200054573-Al.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX
XX 09-APR-1999; 99US-0289346.

```

```

XX  (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI: 2000-618651/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
PT  encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
XX Claim 53; Page 42-43; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
XX
XX Query Match 94.2%; Score 341; DB 21; Length 70;
XX Best Local Similarity 94.3%; Pred. No. 3.5e-35;
XX Matches 66; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVAGRSARGCOTSDNDAAEALNASSKEEALQIRREKIPKYLQFHNLSNL 60
DB 1 tlwgefqvdsrargscqtsndaaalaaasskeaalqirrekipkyllqfhnlnsl 60
QY 61 DRIEDKTPPEP 70
DB 61 drifdktppep 70

RESULT 9
AAB18686
ID AAB18686 standard; peptide: 70 AA.
XX
XX AAB18686:
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX
XX MISC-difference 24 /note= "wild type residue replaced with Leu"
XX
XX MISC-difference 25 /note= "wild type residue replaced with Leu"
XX
XX MISC-difference 26 /note= "wild type residue replaced with Leu"
XX
XX W0200054573-Al.
XX
XX 21-SEP-2000.

```

PF 15-MAR-2000; 2000MO-US06759.  
XX  
PR 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
XX WPI; 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
XX comprise a nucleic acid construct containing a nucleic acid sequence  
XX encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
XX  
XX Claim 53; Page 46; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded  
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1  
XX protein are used to produce transgenic plants. The mutation in AL1 is  
XX present in a ribosome binding region, and expression of mutant AL1  
XX protein imparts increased resistance to geminivirus infection in the  
XX plant. Mutant AL1 proteins are useful for producing plants having  
XX increased resistance or reduced sensitivity to a geminivirus such as  
XX tomo golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
XX virus, cotton leaf curl virus or beet curly top virus.  
SQ Sequence 70 AA;

Query Match 94.2%; Score 341; DB 21; Length 70;  
Best Local Similarity 94.3%; Pred. No. 3.5e-35;  
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLWGEFQVAGRSARGCGCTSDNAAAEALNASSKEEALQIRKIPKYLPOFHNINSL 60  
Db 1 tlwgefygdgrsargcgctsdn11lea1nasskeeaqlirkipekylfqfhninsnl 60

QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

RESULT 10  
AAB18689  
ID AAB18689 standard; peptide; 70 AA.  
XX  
XX AAB18689;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
XX OS Tomato golden mosaic virus.  
XX  
XX  
XX Key Location/Qualifiers  
XX FH Misc-difference 22  
XX FT /note= "wild type residue replaced with Ala"  
XX FT Misc-difference 23  
XX FT /note= "wild type residue replaced with Ala"  
XX PN WO200054573-A1.  
XX

PD 21-SEP-2000.  
XX  
XX  
XX 15-MAR-2000; 2000MO-US06759.  
XX  
XX  
XX 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
XX WPI; 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
XX comprise a nucleic acid construct containing a nucleic acid sequence  
XX encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
XX  
XX Disclosure; Page 48-49; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded  
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1  
XX protein are used to produce transgenic plants. The mutation in AL1 is  
XX present in a ribosome binding region, and expression of mutant AL1  
XX protein imparts increased resistance to geminivirus infection in the  
XX plant. Mutant AL1 proteins are useful for producing plants having  
XX increased resistance or reduced sensitivity to a geminivirus such as  
XX tomo golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
XX virus, cotton leaf curl virus or beet curly top virus.  
SQ Sequence 70 AA;

Query Match 93.9%; Score 340; DB 21; Length 70;  
Best Local Similarity 95.7%; Pred. No. 4.7e-35;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLWGEFQVAGRSARGCGCTSDNAAAEALNASSKEEALQIRKIPKYLPOFHNINSL 60  
Db 1 tlwgefygdgrsargcgctsa1aa1nasskeeaqlirkipekylfqfhninsnl 60

QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

RESULT 11  
AAB18680  
ID AAB18680 standard; peptide; 70 AA.  
XX  
XX AAB18680;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
XX ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
XX OS Tomato golden mosaic virus.  
XX  
XX  
XX Key Location/Qualifiers  
XX FH Misc-difference 42  
XX FT /note= "wild type residue replaced with Ala"  
XX FT Misc-difference 43  
XX FT /note= "wild type residue replaced with Ala"  
XX FT Misc-difference 44

```

FT      /note= "wild type residue replaced with Ala"
XX
XX      WO200054573-A1.
XX
XX      21-SEP-2000.
XX
XX      15-MAR-2000; 2000WO-US06759.
XX
XX      18-MAR-1999; 99US-0125004.
XX      09-APR-1999; 99US-0289346.
XX
XX      (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX      Hanley-Bowdoin L, Orozco BM, Kong L;
XX      WPI: 2000-618851/59.
XX
XX      Transgenic plants with increased resistance to geminivirus infection
XX      comprise a nucleic acid construct containing a nucleic acid sequence
XX      encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX      Claim 52; Page 43-44; 73pp; English.
XX
XX      The present sequence represents a mutant peptide, derived from a
XX      geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX      double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX      DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX      protein are used to produce transgenic plants. The mutation in AL1 is
XX      present in a ribosome binding region, and expression of mutant AL1
XX      protein imparts increased resistance to geminivirus infection in the
XX      plant. Mutant AL1 proteins are useful for producing plants having
XX      increased resistance or reduced sensitivity to a geminivirus such as
XX      tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX      virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX      cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX      virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX      virus, cotton leaf curl virus or beet curly top virus.
XX
XX      Sequence 70 AA:
XX
XX      Query Match          93.4%; Score 338; DB 21; Length 70;
XX      Best Local Similarity 94.3%; Pred. No. 8.3e-35;
XX      Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX      QY      1 TLWGEFOVAGRSARGCGCTSDNAAEALNASSKEEALQIRREKIPKYLFOFHNLSNL 60
XX      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
XX      Db      1 tlwgefvgvdgrsargcgctsdnaaaalnasskeeaqlirrekiptkylfghlnsnl 60
XX
XX      QY      61 DRIFDKTPEP 70
XX      |||||||
XX      Db      61 drifdktpcp 70
XX
XX      RESULT 12
XX      AAB18691
XX      ID      AAB18691 standard; peptide: 70 AA.
XX
XX      AC      AAB18691;
XX      DT      22-JAN-2001 (first entry)
XX
XX      DE      Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX      KW      Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX      ribosome binding region; resistance; geminivirus infection.
XX
XX      OS      Synthetic.
XX      Tomato golden mosaic virus.
XX
XX      Key      Location/Qualifiers
XX      MISC-difference 34

```

```

FT      /note= "wild type residue replaced with Ala"
XX      MISC-difference 35
XX
XX      FT      /note= "wild type residue replaced with Ala"
XX      MISC-difference 36
XX
XX      FT      /note= "wild type residue replaced with Ala"
XX      MISC-difference 36
XX
XX      WO200054573-A1.
XX
XX      21-SEP-2000.
XX
XX      15-MAR-2000; 2000WO-US06759.
XX
XX      18-MAR-1999; 99US-0125004.
XX      09-APR-1999; 99US-0289346.
XX
XX      (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX      Hanley-Bowdoin L, Orozco BM, Kong L;
XX      WPI: 2000-618851/59.
XX
XX      Transgenic plants with increased resistance to geminivirus infection
XX      comprise a nucleic acid construct containing a nucleic acid sequence
XX      encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX      Disclosure: Page 49; 73pp; English.
XX
XX      The present sequence represents a mutant peptide, derived from a
XX      geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX      double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX      DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX      protein are used to produce transgenic plants. The mutation in AL1 is
XX      present in a ribosome binding region, and expression of mutant AL1
XX      protein imparts increased resistance to geminivirus infection in the
XX      plant. Mutant AL1 proteins are useful for producing plants having
XX      increased resistance or reduced sensitivity to a geminivirus such as
XX      tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX      virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX      cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX      virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX      virus, cotton leaf curl virus or beet curly top virus.
XX
XX      Sequence 70 AA:
XX
XX      Query Match          93.4%; Score 338; DB 21; Length 70;
XX      Best Local Similarity 94.3%; Pred. No. 8.3e-35;
XX      Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX      QY      1 TLWGEFOVAGRSARGCGCTSDNAAEALNASSKEEALQIRREKIPKYLFOFHNLSNL 60
XX      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
XX      Db      1 tlwgefvgvdgrsargcgctsdnaaaalnassaaalqirrekiptkylfghlnsnl 60
XX
XX      QY      61 DRIFDKTPEP 70
XX      |||||||
XX      Db      61 drifdktpcp 70
XX
XX      RESULT 13
XX      AAB18681
XX      ID      AAB18681 standard; peptide: 70 AA.
XX
XX      AC      AAB18681;
XX      DT      22-JAN-2001 (first entry)
XX
XX      DE      Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX      KW      Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX      ribosome binding region; resistance; geminivirus infection.
XX
XX      OS      Synthetic.

```

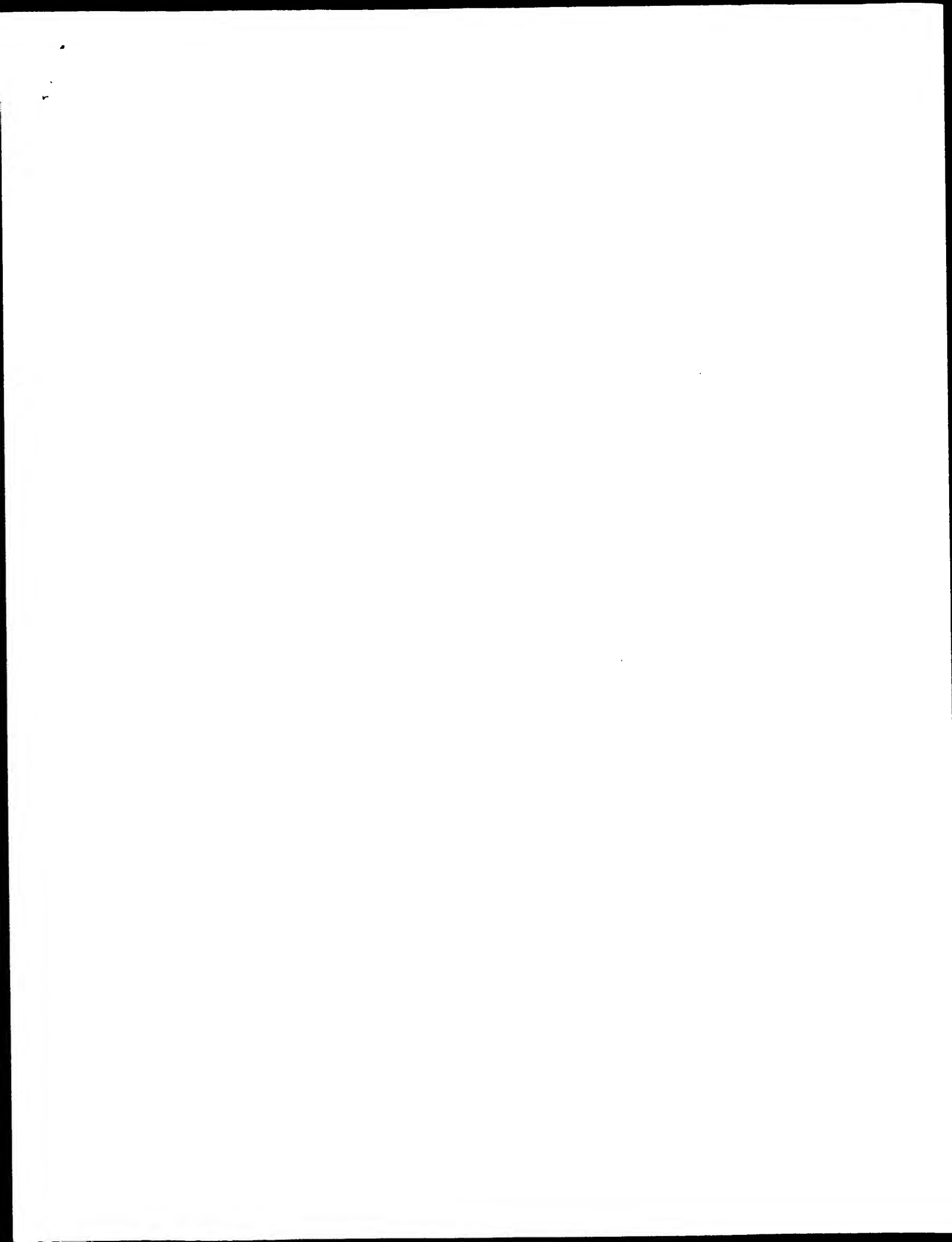


Job time: 528 sec

DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 52 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 54 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
 XX  
 PN W0200054573-AL.  
 XX  
 PD 21-SEP-2000.  
 PD  
 PF 15-MAR-2000; 2000MO-US06759.  
 PF  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI  
 XX WPI: 2000-618651/59.  
 DR  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 53; Page 44-45; 73pp; English.  
 CC  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX  
 SQ Sequence 70 AA:

Query Match 91.7%; Score 332; DB 21; Length 70;  
 Best Local Similarity 94.3%; Pred. No. 4.7e-34;  
 Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVAGRSARGCOTSDNDAAEALNASSKEALQITREKIPKYEYLFQPHNLNSNL 60  
 Db 1 TLVWGEFQVAGRSARGCOTSDNDAAEALNASSKEALQITREKIPKYEYLFQPHNLNSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 DRIFDKTPEP 70

Search completed: October 10, 2002, 01:47:19



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:02:00 ; Search time 118.38 Seconds  
(without alignments)  
102.295 Million cell updates/sec

Title: US-09-289-346a-10  
Perfect score: 364  
Sequence: 1 TLWGEFQVDRSARGCQOT.....FOFHNLSNLDRIKPTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.podent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	300	82.4	351	12	Q91R10
2	298	81.9	352	12	Q9E000
3	281	77.2	226	12	009727
4	281	77.2	226	12	Q9WHF6
5	281	77.2	361	12	067574
6	279	76.6	225	12	090DB1
7	277	76.1	314	12	Q9ELT8
8	277	76.1	364	12	Q9G555
9	274	75.3	185	12	Q98593
10	266	73.1	149	12	P88975
11	266	73.1	233	12	Q9YLA4
12	261	71.7	190	12	Q9Z089
13	261	71.7	190	12	Q9Z084
14	260	71.4	190	12	Q9W827
15	260	71.4	208	12	Q9Z0C4
16	260	71.4	234	12	Q39180

17	260	71.4	359	12	Q91M88
18	258	70.9	208	12	Q9Z0C0
19	258	70.9	208	12	Q9Z0B8
20	256	70.3	208	12	Q9Z0B6
21	255	70.1	203	12	Q9Z0B3
22	251	69.0	363	12	Q73577
23	250	68.7	363	12	Q72719
24	250	68.7	363	12	Q72719
25	249	68.4	349	12	Q88888
26	248	68.1	208	12	Q9Z0A0
27	248	68.1	348	12	Q91M55
28	247	67.9	190	12	Q9Z0B6
29	247	67.9	208	12	Q9Z0C6
30	247	67.9	359	12	Q9Y2V4
31	247	67.9	359	12	Q9Y2V2
32	247	67.9	359	12	Q9Y0X7
33	247	67.9	359	12	Q9YL27
34	246	67.6	359	12	Q88942
35	246	67.6	359	12	Q91B86
36	246	67.6	360	12	Q9DX10
37	245	67.3	362	12	Q91M48
38	244	67.0	307	12	Q91E77
39	244	67.0	361	12	Q72723
40	243	66.8	190	12	Q9Z0A7
41	243	66.8	354	12	Q91M42
42	242	66.5	358	12	Q9JEA2
43	242	66.5	362	12	Q56816
44	241	66.2	231	12	Q96620
45	240	65.9	353	12	Q72692

## ALIGNMENTS

RESULT 1  
ID Q91R10 PRELIMINARY; PRT; 351 AA.  
AC Q91R10;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
RN RPP PROTEIN.  
GN AC1.  
OS Tomato severe rugose virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=158463;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MINAS GERALS;  
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.:  
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting  
RT begomovirus, tomato severe rugose virus, in Brazil."  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY029750; AAK50357.1; -  
SQ SEQUENCE 351 AA: 40122 MW: 87F937A4F873B6CF CRC64;

Query Match 82.4%; Score 300; DB 12; Length 351;  
Best Local Similarity 78.6%; Pred. No. 2.5e-26;  
Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TLWGEFQVDRSARGCQOTNDLLEALNASSKEBALQIIRKIPKYLFOFHNLSNL 60  
DB 111 TIEWGEFQVDRSARGCQOTANDAAALNAPSKDVALQIIRKLPKYLFOFHNLSNL 170  
QY 61 DRIFDKTPEP 70  
DB 171 DRIFARAPEP 180  
RESULT 2  
Q9E000 PRELIMINARY; PRT; 352 AA.





RT hybridization probes and aspects of bean golden mosaic in Brazil.,"  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
 RT "Genetic diversity in geminiviruses causing bean golden mosaic  
 RT disease: The nucleotide sequence of the infectious cloned DNA  
 RT components of a Brazilian isolate of bean golden mosaic geminivirus.,"  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M88686; AAA46312.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRODOM: PD000736; Gemini\_AL1.  
 DR SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.28; Score 281; DB 12; Length 361;  
 Best Local Similarity 77.6%; Pred. No. 3.9e-24;  
 Matches 52; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 4 WGEFVDGSRARGGCGTNDLLLEALNASSKEPALQITREKIPKYLFOFHNLNSNDRI 63  
 Db 113 WGHFVDGSRARGGCGTNDAAEALNASSKEPALQITREKIPKYLFOYHNLSSNDRI 172  
 OY 64 FDKTPPEP 70  
 Db 173 FTKAPDP 179

RESULT 6  
 Q9QDB1 PRELIMINARY; PRT; 225 AA.

AC 09QDB1;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS cowpea golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=69263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGMV-BR;  
 RA Faria J.C.;  
 RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from  
 RT Brazil.,"  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF188708; AAF06318.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRODOM: PD000736; Gemini\_AL1.  
 DR NON\_TER 225  
 FT SEQUENCE 225 AA; 25766 MW; 1089C6B8D8D15B5D CRC64;

Query Match 76.68; Score 279; DB 12; Length 225;  
 Best Local Similarity 77.6%; Pred. No. 3.8e-24;  
 Matches 52; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 4 WGEFVDGSRARGGCGTNDLLLEALNASSKEPALQITREKIPKYLFOFHNLNSNDRI 63  
 Db 113 WGHFVDGSRARGGCGTNDAAEALNASSKEPALQITREKIPKYLFOYHNLSSNDRI 172  
 OY 64 FDKTPPEP 70  
 Db 173 FTKAPDP 179

RESULT 7  
 Q9ELT8

ID Q9ELT8 PRELIMINARY; PRT; 314 AA.  
 AC Q9ELT8;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION ASSOCIATION PROTEIN.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 RT States.,"  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A. to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF288227; AAG01006.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRODOM: PD000736; Gemini\_AL1.  
 DR SEQUENCE 314 AA; 35153 MW; 686220613046943F CRC64;

Query Match 76.1%; Score 277; DB 12; Length 314;  
 Best Local Similarity 79.4%; Pred. No. 9.5e-24;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 1 TLVWGFVDGSRARGGCGTNDLLLEALNASSKEPALQITREKIPKYLFOFHNLNSNL 60  
 Db 110 TLVWGFVDGSRARGGCGTNDAAEALNASSKEPALQITREKIPKYLFOYHNLVSNL 169  
 OY 61 DRIFDKTP 68  
 Db 170 DRIFSPPP 177

RESULT 8  
 Q9QS55 PRELIMINARY; PRT; 364 AA.

ID Q9QS55;  
 AC 09QS55;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 GN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De la Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 RT States.,"  
 RL Plant Dis. 82:1253-1257(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A.;  
 RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl  
 RT virus : nucleotide sequence and phylogenetic relationships.,"  
 RL Online Publication.  
 DR EMBL: AF104036; AAD47173.1; -;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRODOM: PD000736; Gemini\_AL1.  
 DR SEQUENCE 364 AA; 40680 MW; 5F79752A31A09D6E CRC64;



```

ID Q92089 PRELIMINARY: PRT: 190 AA.
AC Q92089;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YOKOHAMA3;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL, AB001315; BAA34033.1; -
DR InterPro: IPR001191; Geminl_AL1.
DR Pfam: PF00799; Geminl_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Geminl_AL1.
FT NON_TER 1
FT SEQUENCE 190 AA: 21432 MW; AAC093D1D1610FAD CRC64;
SQ

```

```

Query Match 71.7%; Score 261; DB 12; Length 190;
Best Local Similarity 60.0%; Pred. No. 3.7e-22;
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

```

```

QY 1 TLVWGEFVDGRSARGCGCOTSDNLLLEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60
DB 85 TLEWGTFOIDGRSARGCGQNDACAELNASSKAELALIREKLPKFIQYHNLSNL 144
QY 61 DRI-----FDKTP 69
DB 145 DRIFAPPLEVFCPTASSFDVPE 169

```

```

RESULT 13
ID Q92084 PRELIMINARY: PRT: 190 AA.
AC Q92084;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-2, PARTIAL
DE AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YOKOHAMA5;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL, AB001318; BAA34039.1; -
DR InterPro: IPR001191; Geminl_AL1.
DR Pfam: PF00799; Geminl_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Geminl_AL1.
FT NON_TER 1
FT SEQUENCE 190 AA: 21444 MW; AAC1C2943E3F01AD CRC64;
SQ

```

```

Query Match 71.7%; Score 261; DB 12; Length 190;
Best Local Similarity 60.0%; Pred. No. 3.7e-22;
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

```

```

QY 1 TLVWGEFVDGRSARGCGCOTSDNLLLEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60
DB 85 TLEWGTFOIDGRSARGCGQNDACAELNASSKAELALIREKLPKFIQYHNLSNL 144
QY 61 DRI-----FDKTP 69
DB 145 DRIFAPPLEVFCPTASSFDVPE 169

```

```

RESULT 14
ID Q9W827 PRELIMINARY: PRT: 190 AA.
AC Q9W827;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORAL-1, PARTIAL AND
DE COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GORAL;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL, AB001303; BAA34010.1; -
DR InterPro: IPR001191; Geminl_AL1.
DR Pfam: PF00799; Geminl_AL1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Geminl_AL1.
FT NON_TER 1
FT SEQUENCE 190 AA: 21444 MW; 93C3742A8EBD87EB CRC64;
SQ

```

```

Query Match 71.4%; Score 260; DB 12; Length 190;
Best Local Similarity 60.0%; Pred. No. 4.8e-22;
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

```

```

QY 1 TLVWGEFVDGRSARGCGCOTSDNLLLEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60
DB 85 TLEWGTFOIDGRSARGCGQNDACAELNASSKAELALIREKLPKFIQYHNLSNL 144
QY 61 DRI-----FDKTP 69
DB 145 DRIFAPPLEVFCPTASSFDVPE 169

```

```

RESULT 15
ID Q9Z0C4 PRELIMINARY: PRT: 208 AA.
AC Q9Z0C4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL AND
DE COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ABURA3;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL, AB001294; BAA33992.1; -
DR InterPro: IPR001191; Geminl_AL1.
DR Pfam: PF00799; Geminl_AL1.

```

DR PRINTS: PR00227; GEMCOATALL.  
 DR Prodom: PD000736; Gemini\_ALI: 1.  
 FT NON\_TER 1  
 FT NON\_TER 208  
 SQ SEQUENCE 208 AA: 23526 MW: 249CC31D8729C72D CRC64:

Query Match 71.4%; Score 260; DB 12; Length 208;  
 Best Local Similarity 60.0%; Pred. No. 5.3e-22;  
 Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

OY 1 TLVWGEFQVGRSRGGCCOTSNLLEALINASSKEEALQIIRKIPKYLFOFHNLNSNL 60  
 Db 102 TLEWGTFOYDGRSARGGCCQANDACAEALNASSKAELSTIREKLPKDFIFQYHNLNSNL 161

OY 61 DRI-----FDKTPK 69  
 Db 162 DRIFAPPLEVFCPPFSNSFDQVPE 186

Search completed: October 10, 2002, 02:02:00  
 Job time: 1093 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:53 ; Search time 34.62 Seconds  
(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346a-10  
Perfect score: 364  
Sequence: 1 TLWGCEFOVDGRSARGCQT.....FOFHNLNSLDRIFDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	95.9	352	1 VAL1_TGMV	P03567 tomato gold
2	249	68.4	361	1 VAL1_PYMV	P27258 potato yell
3	242	66.5	358	1 VAL1_CLVK	P14982 cassava lat
4	242	66.5	358	1 VAL1_CLVN	P14972 cassava lat
5	235	64.6	362	1 VAL1_TYLC	P36279 tomato yell
6	228	62.6	349	1 VAL1_PHV	P06923 pepper huas
7	228	62.6	359	1 VAL1_TYLCU	P38609 tomato yell
8	221	60.7	353	1 VAL1_BGMV	P05175 bean golden
9	221	60.7	359	1 VAL1_TYLCM	P27260 tomato yell
10	220	59.4	355	1 VAL1_ABMVW	P21947 beet curly
11	217	59.6	358	1 VAL1_BCTV	P14991 beet curly
12	217	59.6	361	1 VAL1_TMOV	P06657 tomato yell
13	206	56.6	357	1 VAL1_TYLCV	P27259 tomato yell
14	125	34.3	347	1 VAL1_STCV	P29048 squash leaf
15	63	17.3	342	1 RSMC_ECCL	P39406 escherichia
16	62	17.0	295	1 VAL1_TYDVA	P31617 tobacco yell
17	62	17.0	387	1 Y4PF_RHISN	P55615 bacillus su
18	61	16.8	555	1 GLPD_BACSU	P18158 bacillus su
19	61	16.8	836	1 GCSR_HUMAN	P09062 homo sapien
20	60.5	16.6	447	1 TBL2_HUMAN	P09443 homo sapien
21	60	16.5	470	1 RHBH_RHIME	P09233 rhizobium m
22	60	16.5	509	1 C4A3_DROME	P09471 streptomyce
23	60	16.5	799	1 AFSK_STRCO	P54742 streptomyce
24	60	16.5	807	1 AFSK_STRCR	P54742 streptomyce
25	60	16.5	1117	1 TEPT_TERTH	P77448 tetrahymena
26	59.5	16.3	79	1 AKHD_DROME	P17975 drosophila
27	58.5	16.1	323	1 VAL1_PASV	P00038 paniscus str
28	58.5	16.1	397	1 METL_RAT	P13444 ratius norv
29	58.5	16.1	1610	1 CCAD_MESAU	P09244 mesocricetu
30	58.5	16.1	2161	1 CCAD_HUMAN	P01668 homo sapien
31	58.5	16.1	2203	1 CCAD_RAT	P27732 ratius norv
32	58	15.9	256	1 YKJ9_YEAST	P34247 saccharomyc
33	58	15.9	354	1 CARA_METJA	P58425 methanococ

34	58	15.9	543	1 NFL_HUMAN	P07196 homo sapien
35	58	15.9	1287	1 SKI2_YEAST	P35207 saccharomyc
36	57	15.7	784	1 LON_RICPR	P09292 rickettsia
37	57	15.7	970	1 DAB1_YEAST	P21657 saccharomyc
38	57	15.7	1517	1 RPOC_CAMUE	P09130 campylobact
39	56.5	15.5	266	1 ETPB_MYCUE	P03095 mycobacteri
40	56.5	15.5	266	1 ETPB_MYCUE	P03095 mycobacteri
41	56.5	15.5	395	1 METL_HUMAN	P00266 homo sapien
42	56.5	15.5	1026	1 MYTB_DROME	P02397 drosophila
43	56.5	15.5	1042	1 TIRH_DROME	P060295 methanococ
44	56	15.4	247	1 YCP4_YEAST	P25349 saccharomyc
45	56	15.4	299	1 Y175_HELPY	P56112 helicobacte

## ALIGNMENTS

RESULT 1	VAL1_TGMV	STANDARD:	PRT: 352 AA.
ID	VAL1_TGMV		
AC	P03567:		
DT	21-JUL-1986 (rel. 01, Created)		
DT	21-JUL-1986 (rel. 01, Last sequence update)		
DT	01-JUN-1994 (rel. 29, Last annotation update)		
DE	ALI protein.		
GN	AC1.		
OS	Tomato golden mosaic virus (TGMV).		
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.		
ON	NCBI_Taxid=10831;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck R.W.;		
RT	*Complete nucleotide sequence of the infectious cloned DNA components		
RT	of tomato golden mosaic virus: potential coding regions and regulatory		
RT	sequences.		
RL	EMBO J. 3:2197-2205(1984).		
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> )		
CC	-----		
DR	EMBL; K02029; -; NOT_ANNOTATED_CDS.		
DR	PIR; A04170; QOCVLI.		
DR	InterPro; IPR001191; Geminl_ALI.		
DR	Pfam; PF00799; Geminl_ALI; 1.		
DR	PRINTS; PR00227; GEMCOATALI.		
DR	PRODOM; PD000736; Geminl_ALI; 1.		
KW	ATP-binding.		
FT	NP-BIND 223		
FT	SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;		
QY	1 TLWGCEFOVDGRSARGCQTSNDLLLEALNASKKEALDIIREKIPEKTLFOFHNLNSLU 60		
DB	111 TLWGCEFOVDGRSARGCQTSNDAAAEALNASKKEALDIIREKIPEKTLFOFHNLNSLU 170		
QY	61 DRIFDKTPPEP 70		
DB	171 DRIFDKTPPEP 180		
RESULT 2	VAL1_PYMV	STANDARD:	PRT: 361 AA.
ID	VAL1_PYMV		

AC P27258;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE AL1 protein.  
 OS Potato yellow mosaic virus (isolate Venezuela).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10828;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91311403; PubMed=1856690;  
 RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;  
 RT "The nucleotide sequence of the infectious cloned DNA components of  
 RT potato yellow mosaic virus."  
 RL J. Gen. Virol. 72:1515-1520(1991).  
 CC -i- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D00940; BAA00782.1; -.  
 DR PIR: J00364; OOCVPT.  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 KM ATP-binding.  
 FT NP\_BIND 222 229 ATP (POTENTIAL).  
 SQ SEQUENCE 361 AA: 40850 MW: 5627A33BF1264383 CRC64:

Query Match 68.4%; Score 249; DB 1; Length 361;  
 Best Local Similarity 66.7%; Pred. No. 2.3e-20;  
 Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVNDGRSARGCGQTSNDLLLEALNASSKEBALQIIREKIPKYLQFPHNLNSNL 60  
 DB 110 TLEWELFQIDGRSARGGQGVNDAALNAGCTKEAMKIKELPEKFLQYHNLNSNL 169  
 OY 61 DRIFDKTPE 69  
 DB 170 DRIFMKAPE 178

RESULT 3  
 VAL1\_CLVK STANDARD; PRT; 358 AA.  
 AC P14982;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein (40.4 kDa protein).  
 GN AC1.  
 OS Cassava latent virus (strain West Kenyan 844).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Gay M.R.;  
 RT "Nucleotide sequence of cassava latent virus DNA."  
 RL Nature 301:260-262(1983).  
 CC -i- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J02057; -; NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 KM ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA: 40346 MW: ED173E753EE92D69 CRC64:

Query Match 66.5%; Score 242; DB 1; Length 358;  
 Best Local Similarity 61.4%; Pred. No. 1.3e-19;  
 Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVNDGRSARGCGQTSNDLLLEALNASSKEBALQIIREKIPKYLQFPHNLNSNL 60  
 DB 109 TVEWGEFQIDGRSARGGQSANDAYAKALNAGSKSEALNVIRELVKDFVLOFPHNLNSNL 168  
 OY 61 DRIFDKTPEP 70  
 DB 169 DRIFPEPPAP 178

RESULT 4  
 VAL1\_CLVK STANDARD; PRT; 358 AA.  
 AC P14972;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 protein (40.4 kDa protein).  
 GN AC1.  
 OS Cassava latent virus (strain Nigerian).  
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10819;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90174930; PubMed=2308831;  
 RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;  
 RT "Nucleotide sequence of the infectious cloned DNA components of  
 RT African cassava mosaic virus (Nigerian strain)."  
 RL Nucleic Acids Res. 18:197-198(1990).  
 CC -i- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X17095; CAA34953.1; -.  
 DR PIR: S07594; S07594.  
 DR InterPro: IPR001191; Geminl\_AL1.  
 DR Pfam: PF00799; Geminl\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATAL1.  
 DR PRODOM: PD000736; Geminl\_AL1; 1.  
 KM ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA: 40435 MW: 1DB16B0CB2DE52C CRC64;

Query Match 66.5%; Score 242; DB 1; Length 358;  
 Best Local Similarity 61.4%; Pred. No. 1.3e-19;  
 Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVNDGRSARGCGQTSNDLLLEALNASSKEBALQIIREKIPKYLQFPHNLNSNL 60  
 DB 109 TVEWGEFQIDGRSARGGQSANDAYAKALNAGSKSEALNVIRELVKDFVLOFPHNLNSNL 168



Best Local Similarity 59.4%; Pred. No. 4.8e-18;  
Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

OY 2 LWGGEFQVDRSARGGCGTNDLLEALNASSKEEALQIIRKIPKYLFPFHNLNSMD 61  
DB 111 LEMGTQIDGRSARGGCGTNDVAKALNAGSKSEALDYKELAPRDYVLFPHNINSMD 170

OY 62 RIFDKTPEP 70  
DB 171 RVEQVPPAP 179

RESULT 8  
VALL\_BGMV STANDARD; PRT; 353 AA.  
ID VALL\_BGMV  
AC P05175;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE AL1 protein (40.2 kDa protein).  
GN AC1.  
OS Bean golden mosaic virus.  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;  
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene  
RT regulation in geminiviruses.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL; M10070; AAA6318.1;  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR ProDom: PD000736; Gemini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 222 229 ATP (POTENTIAL).  
SQ SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match 60.7%; Score 221; DB 1; Length 353;  
Best Local Similarity 60.0%; Pred. No. 2.8e-17;  
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLWGEFQVDRSARGGCGTNDLLEALNASSKEEALQIIRKIPKYLFPFHNLNSMD 60  
DB 110 TIEMGFQVDRSARGGCGTNDVAKALNAGSKSEALDYKELAPRDYVLFPHNINSMD 169

OY 61 RIFDKTPEP 70  
DB 170 ERIFKVPAP 179

RESULT 9  
VALL\_TYLCM STANDARD; PRT; 359 AA.  
ID VALL\_TYLCM  
AC P27260;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 protein (C1 protein).  
GN C1.  
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).

Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=92107660; PubMed=1840676;  
RA Khayr-Pour A., Bendahmane M., Matzelt V., Accotto G.P., Crespi S.,  
RA Gronenborn B.;  
RT "Tomato yellow leaf curl virus from Sardinia is a  
RT whitefly-transmitted monopartite geminivirus.";  
RL Nucleic Acids Res. 19:6763-6769(1991)  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL; X61153; CAA43466.1;  
DR PIR: S22593; S22593.  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR ProDom: PD000736; Gemini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 220 227 ATP (POTENTIAL).  
SQ SEQUENCE 359 AA; 40733 MW; 9717B4A0C93EFA7 CRC64;

Query Match 60.7%; Score 221; DB 1; Length 359;  
Best Local Similarity 56.5%; Pred. No. 2.9e-17;  
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 2 LWGGEFQVDRSARGGCGTNDLLEALNASSKEEALQIIRKIPKYLFPFHNLNSMD 61  
DB 111 LEMGTQIDGRSARGGCGTNDVAKALNAGSKSEALDYKELAPRDYVLFPHNINSMD 170

OY 62 RIFDKTPEP 70  
DB 171 RVEQVPPAP 179

RESULT 10  
VALL\_ABMVW STANDARD; PRT; 355 AA.  
ID VALL\_ABMVW  
AC P21947;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 protein.  
GN AC1.  
OS Abutilon mosaic virus (isolate West India).  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=91020984; PubMed=2219703;  
RA Frischmuth T., Zimmet G., Jeske H.;  
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
RT as well as eukaryotic features.";  
RL Virology 178:461-468(1990).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC



DR EMBL: X15983; -: NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; Q0CWVL.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 DR ATP-binding.  
 KW NP\_BIND.  
 FT NE\_BIND.  
 SQ SEQUENCE 355 AA; 40257 MW; 16A2CABA63251E95 CRC64;

Query Match 60.4%; Score 220; DB 1; Length 355;  
 Best Local Similarity 58.6%; Pred. No. 3.7e-17;  
 Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRSGGCGTSDLLLEALNASSKEEALQIRKIPKXYLFQFHNLSNL 60  
 DB 110 TAEWGEFQIDGRSARGGQOTANDSYAKALNAGVQSALNTLKEOPKDYVLQNHNRSL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 11  
 VAL1\_BCTV STANDARD; PRT: 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE A1 protein (40.8 kDa protein).  
 OS Beet curly top virus (BCTV).  
 CC Viruses: ssDNA viruses; Geminiviridae; Curtovirus.  
 OX NCBI\_TaxID=10840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Plummer M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 RT beet curly top virus."  
 RL EMBO J. 5:1761-1767(1986).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X04144; -: NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 DR ATP-binding.  
 KW NP\_BIND.  
 FT NE\_BIND.  
 SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C33 CRC64;

Query Match 59.6%; Score 217; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 8e-17;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRSGGCGTSDLLLEALNASSKEEALQIRKIPKXYLFQFHNLSNL 60  
 DB 110 TAEWGEFQIDGRSARGGQOTANDSYAKALNAGVQSALNTLKEOPKDYVLQNHNRSL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 OKIFQRPDP 179

RESULT 12  
 VAL1\_TMOV STANDARD; PRT: 361 AA.  
 AC O06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein.  
 GN AL1.  
 OS Tomato mottle virus (isolate Florida) (TMov).  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93107858; PubMed=1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 RT isolated from tomatoes in Florida."  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L14460; AAC32414.1; -.  
 DR PIR: J01870; J01870.  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 DR ATP-binding.  
 KW NP\_BIND.  
 FT NE\_BIND.  
 SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEBAC6950 CRC64;

Query Match 59.6%; Score 217; DB 1; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 8.1e-17;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRSGGCGTSDLLLEALNASSKEEALQIRKIPKXYLFQFHNLSNL 60  
 DB 110 TAEWGEFQIDGRSARGGQOSANDSYAKALNAGVQSALVAREOPKDYVLQNHNRSL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 ERIFAKAPEP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT: 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE A1 protein (C1 protein).  
 GN C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RT with a single genomic component."  
 RL Virology 185:151-161(1991).

```

CC -! SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15656; CAA3688.1; -.
DR PIR: D40779; QOCVC1.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Gemini_AL1.1.
DR ATP-binding.
DR NP_BIND
FT SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;
SQ

```

```

Query Match 56.6%; Score 206; DB 1; Length 357;
Best Local Similarity 63.9%; Pred. No. 1.3e-15;
Matches 39; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

```

```

OY 4 WGEFVQDGRSARGCGQTSNDLLLEALNASKKEALQIREKIPKYPFQPHNLSNDRI 63
DB 111 FVSGQIDGRSARGGQGSANDATAYEALNKGSKSEALNLIKERAKPDYIIQFHNLSNDRI 170
OY 64 F 64
DB 171 F 171

```

## RESULT 14

```

VALL_SLCV STANDARD; PRT; 347 AA.
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE All protein.
OS Squash leaf curl virus.
OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE=91082449; PubMed=1984668;
RA Lazarewitz S.G., Lazdins J.B.;
RT "Infectivity and complete nucleotide sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype."
RL Virology 180:58-69(1991).
CC -! SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M38183; AAC32410.1; ALT_INT.
DR PIR: G36785; QOCVSL.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOATALL.
DR ProDom: PD000736; Gemini_AL1.1.
DR ATP-binding.
DR NP_BIND
FT SEQUENCE 347 AA; 39110 MW; AFDABEDDE122110E CRC64;
SQ

```

```

Query Match 34.3%; Score 125; DB 1; Length 347;
Best Local Similarity 37.9%; Pred. No. 1.3e-06;
Matches 25; Conservative 15; Mismatches 22; Indels 4; Gaps 1;

```

```

OY 5 GEFVQDGRSARGCGQTSNDLLLEALNASKKEALQIREKIPKYPFQPHNLSNDRI 64
DB 116 GYKRVSG---GSKSNKDDYHNAVNAAGSEALDIKKADPKPTFLVYHNLANVERLF 171
OY 65 DKTRPP 70
DB 172 QKPEPP 177

```

## RESULT 15

```

RSMC_ECOLI STANDARD; PRT; 342 AA.
ID RSMC_ECOLI
AC P39406;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribosomal RNA small subunit methyltransferase C (EC 2.1.1.52) (rRNA
DE (guanine-N2-)-methyltransferase) (16S rRNA m2G1207 methyltransferase).
GN RSMC OR B4371.
OS Escherichia coli.
OC Bacteria: Proteobacteria: gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN 11
RP SEQUENCE FROM N.A.
RA STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes."
RL Nucleic Acids Res. 23:2105-2119(1995).
RN 12
RP SEQUENCE OF 1-19, AND CHARACTERIZATION.
RX MEDLINE=99091632; PubMed=9873033;
RA Tscherner J.S., Nurse K., Popienick P., Ofengand J.;
RT "Purification, cloning, and characterization of the 16S RNA m2G1207
RT methyltransferase from Escherichia coli."
RL J. Biol. Chem. 274:924-929(1999).
CC -! FUNCTION: SPECIFICALLY METHYLATES THE GUANOSINE IN POSITION 1207
CC OF 16S RNA IN THE 30S PARTICLE.
CC -! CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing N2-methylguanine.
CC -! COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.
CC -! SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. RSMC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14003; AAA97267.1; -.
DR EMBL: AE000507; AAC77324.1; -.
DR EcoGene: EG12596; rsmc.
DR InterPro: IPR002052; N6_Mtase.
DR InterPro: IPR000051; SAM_bind.
DR PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
DR rRNA processing; Transferase; Methyltransferase; Magnesium;
DR Complete proteome.
DR INIT_MET
FT SEQUENCE 342 AA; 37493 MW; C7A318155700302D CRC64;
SQ

```

```

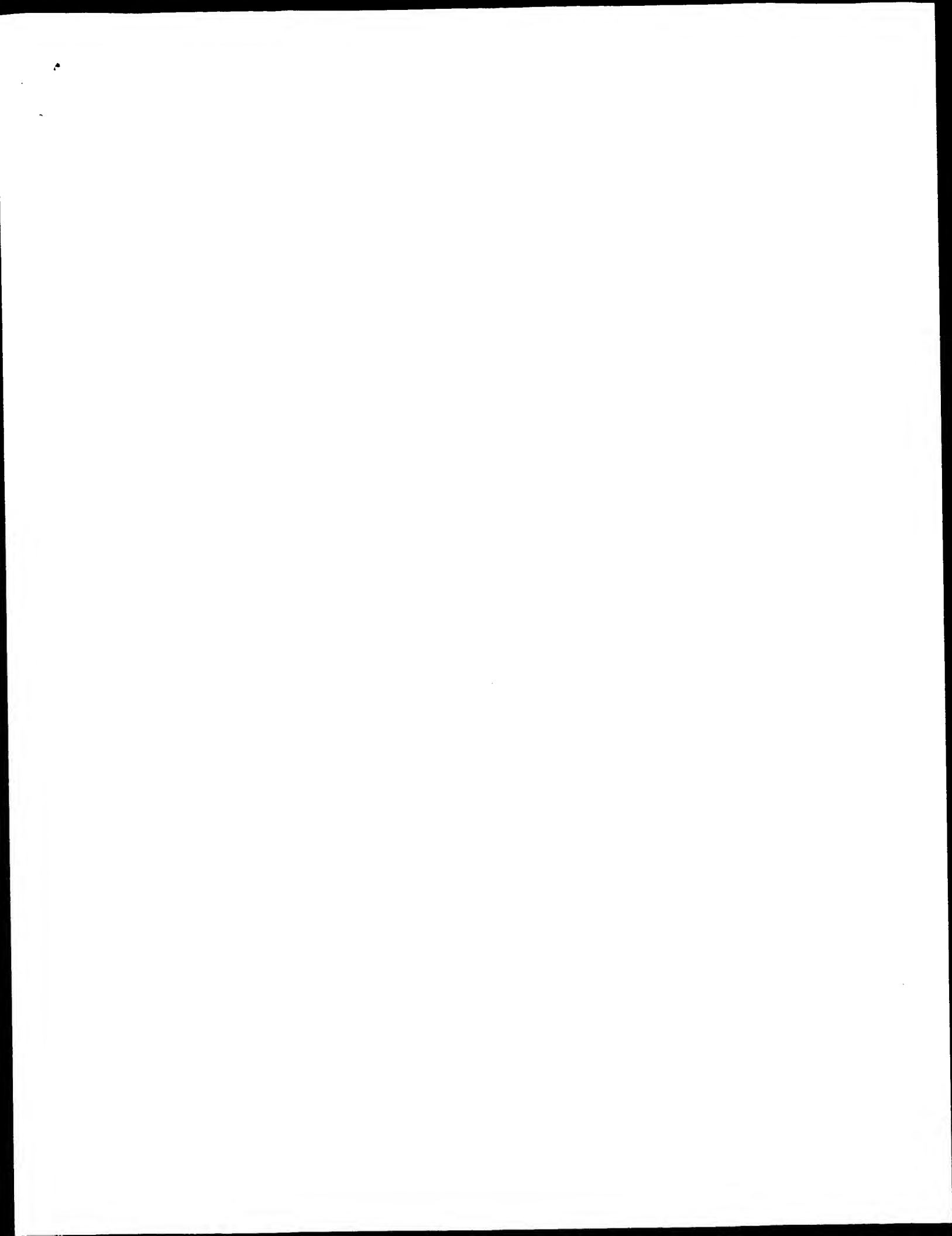
Query Match 17.3%; Score 63; DB 1; Length 342;

```

Best Local Similarity 31.1%; Pred. No. 9.5;  
Matches 14; Conservative 7; Mismatches 16; Indels 8; Gaps 1;

QY 4 WGEFOVDGRS-----ARCGCQTSNDLLPALNASSKKEALQI 40  
|||: |||: : |||: | : | :  
Db 158 WGEYSVDGLTVKTLPGVFSRSDGLDVGSQLLLSTLTPTHTKGKVLDV 202

Search completed: October 10, 2002, 02:09:53  
Job time: 1351 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:45 ; Search time 69.63 seconds  
(without alignments) 96.600 Million cell updates/sec

Title: US-09-289-346a-10

Sequence: 1 TLVWGEFQVDCRSARGCOT.....FOFHNLNSLDRIDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR-71:\*  
2: pir1:\*  
3: pir2:\*  
4: pir3:\*  
5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	349	95.9	352	1	00CVL1	Al1 protein - toma
2	249	68.4	361	1	00CVPT	Al1 protein - pola
3	242	66.5	358	2	S07594	hypothetical prote
4	235	64.6	362	1	J01887	Al1 protein - toma
5	228	62.6	349	2	J02300	replicase - pepper
6	228	62.6	349	2	S31875	Al1 protein - pepp
7	228	62.6	359	2	S39211	gene C1 protein -
8	221	60.7	359	2	S22593	hypothetical prote
9	220	60.4	355	1	00CVY1	AV1 protein - abut
10	218	59.9	351	2	J02327	Al1 protein - inda
11	217	59.6	358	1	J01870	Al1 protein - toma
12	217	59.6	385	2	S28360	Al1 protein - beet
13	215	59.1	359	2	S39235	gene C1 protein -
14	208	57.1	360	2	S59885	replication-associ
15	206	56.6	357	1	00CVCI	Al1 protein - toma
16	138	37.9	347	1	S45059	AC1 protein (clone
17	125	34.3	347	1	00CVS1	Al1 protein (squa
18	67	18.4	714	2	C95382	probable ferrichto
19	64	17.6	154	2	AD3475	nitrogen regulator
20	64	17.6	587	2	JC1419	Fc gamma (19c) rec
21	63.5	17.4	1229	2	T48859	kinasin-like prote
22	63	17.3	343	2	A98295	ribosomal RNA smal
23	63	17.3	343	2	C86136	probable enzyme yj
24	63	17.3	343	2	S56595	rRNA (guanine-N2-)
25	62	17.0	295	2	D42452	Cl protein - tobac
26	62	17.0	329	2	D85056	hypothetical prote
27	62	17.0	329	2	A91111	hypothetical prote
28	62	17.0	416	2	A82892	hypothetical prote
29	61.5	16.9	333	2	B42476	hypothetical prote

30	61.5	16.9	333	2	AC0766	probable glycosylt
31	61.5	16.9	481	2	A70091	probable phosphos
32	61.5	16.9	1279	2	E64709	type IIS restricti
33	61	16.8	555	2	C45868	glycerol-3-phospha
34	61	16.8	771	2	B38252	granulocyte colony
35	61	16.8	783	2	JH0329	granulocyte colony
36	61	16.8	863	2	C38252	granulocyte colony
37	60.5	16.6	447	2	T12544	granulocyte prote
38	60.5	16.6	1792	2	T13939	myosin V - fruit f
39	60	16.5	247	2	H86844	glutamine ABC tran
40	60	16.5	470	2	T46814	gamma-aminobutyrat
41	60	16.5	470	2	B95419	diaminobutyrate--p
42	60	16.5	799	2	T48889	serine/threonine P
43	60	16.5	1117	2	T14891	telomerase (EC 2.7
44	60	16.5	1265	2	T02131	hypothetical prote
45	59.5	16.3	79	1	A58656	adipokinetic hormo

#### ALIGNMENTS

RESULT 1  
00CVL1  
AL1 protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C>Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
C:Accession: A04170  
R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBO J. 3, 2197-2205, 1984  
A>Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAN>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.9%; Score 349; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 2.6e-31;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDCRSARGCCTSDLLLEALNASSKEBALDIIEKIPKYLFOFHNLNSNL 60  
DB 111 TLVWGEFQVDCRSARGCCTSDNMAAEALNASSKEBALDIIEKIPKYLFOFHNLNSNL 170

QY 61 DRIFDKTPPEP 70  
DB 171 DRIFDKTPPEP 180

RESULT 2  
00CVPT  
AL1 protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
J: Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A>Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; M01D:91311403  
A:Accession: J00364  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-361 <CON>  
C:Cross-references: GB:ID00940; NID:9222458; PIDN:BA00782.1; PID:9222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein



Query Match	59.6%;	Score 217;	DB 1;	Length 358;
Best Local Similarity	55.7%;	Pred. No. 1.5e-16;		
Matches 39;	Conservative 15;	Mismatches 16;	Indels 0;	Gaps 0;

```

OY      1 TLVGEFVQDGRSARGCGQTINDLLLEALNNSSKEEALQIIRKIPKTYLFQFHNLSNL 60
      :|::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      107 TIEMDDFOIDGRSARGGQASANDSYAKALNAGVSQALVLRKQPKDFVLQNNIRSNL 166
OY      61 DRIFDKTREP 70
      :|||::|||::|||
Db      167 ERIFAKAPRP 176

RESULT 12
S28360
All protein - beet curly top virus
C:Species: beet curly top virus
C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C:Accession: S28360
R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A>Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top
A:Reference number: S28360
A:Accession: S28360
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-385 <STA>
A:Cross-references: GB:M4597; EMBL:X04144; NID:9210678; PIDN:AAA42751.1; PID:9210679
C:Superfamily: tomato golden mosaic virus All protein

```

	Query Match	59.6%	Score 217;	DB 2:	Length 385;	
	Best Local Similarity	55.7%	Pred.	No. 1.7e-16;		
	Matches	39;	Conservative	15;	Mismatches	16;
					Indels	0;
					Gaps	0;
QY	1	TIVWGEFVNGRARGAGCOTSDNDLLEALMASKKEALITIREKIPEKYLFQFHNLNSL	60			
		:         :               :     :         :     :         :				
Dd	137	TIEGEEFOIDGRSARSGOOTANDSTAKALNATSLDQALIKEDPKDYFLQHNNLNNA	196			
QY	61	DRIPDKTPRP	70			
		:   :  :   :				
Dd	197	OKITFORPPDP	206			

RESULT 13  
S39235  
gene C1 protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
C:Accession: S39235  
R:Crespi, S.; Morris, E.; Vaira, A.; Bosco, D.; Accotto, G.  
submitted to the EMBL Data Library, December 1993  
A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity  
A:Reference number: S39233  
A:Accession: S39235  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-359 <CRE>  
A:Cross-references: EMBL:Z28390; NID:q1041671; PID:q1334964  
C:Superfamily: tomato golden mosaic virus AL1 protein

	Query Match	59.1%;	Score 215;	DB 2;	Length 359;	
	Best Local Similarity	55.1%;	Pred. No.	2.5e-16;		
	Matches	38;	Conservative	12;	Mismatches	19; Indels
					0;	Gaps
OY	2	LWGEFOVGRSARGGCGTNDLLEALNASSKEEALQIREKPEKYDFOFHNLNSDLD	61			
	:                   :	:                 :	:			
Dd	111	LWMGTGQIDGRSARGGOQTANDATAKAINARKSEFLADVIKQLAPRDYVLHFHNISNDL	170			
OY	62	RIFDKTPRP	70			
	:::					
Dd	171	KVFQVPAP	179			

RESULT 14

S59885  
replication-associated protein C1 - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
C:Accession: S59885  
R:Hong, Y.; Harrison, B.D.  
Submitted to the EMBL Data Library, February 1995  
A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
d geminiviruses.  
A:Reference number: S58346  
A:Accession: S59885  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <HON>  
A:Cross-references: EMBL:Z48182; NID:9444838; PIDN:CMA88229.1; PID:9974211  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	57.1%;	Score 208;	DB 2;	Length 360;
Best Local Similarity	59.1%;	Pred. No. 1,66-15;		
Matches	39;	Conservative 12;	Mismatches 15;	Indels 0;
Gaps	0;			
QY	4	WGEQVQVRSARGCGCTSDNDLLLEALNASSKEALQIIREKIPERYLFFQFNILNSLDRI	63	
	:::::	:::::	:::	
DB	113	FGVAFQIDGRSARGGQGSANDAYAEALNSSKAAALDIIRKAPDFVLQFNINLANLDRI	172	
QY	64	FDKTPPE	69	
	:::			
DB	173	FTPSAE	178	

000VC1  
 ALL protein - tomato yellow leaf curl virus  
 N:Alternate names: CI protein  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: DA0779  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 189, 151-161, 1991  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin-  
 A:Reference number: AA0779; MUID:92024070  
 A:Accession: DA0779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NA>  
 A:Cross-references: GB:X15656; NID:G62204; PIDN:CAA33688.1; PID:G62207  
 C:Superfamily: tomato golden mosaic virus ALL protein

	Query Match	56.6%;	Score 206;	DB 1;	Length 357;	
	Best Local Similarity	63.9%;	Pred. No.	2.6e-15;		
	Matches	39;	Conservative	10;	Mismatches 12;	Indels 0; Gaps 0.
OY	4 WGEFQVNGRSARGGCOTSDNLLEALNASSKEEALQIREKIPERYLFOPHNLNSNDRI	63	:	:	: ::   ::   :	: ::   ::   :
Dd	111 FGVSQIDGRSARGGGSANDAYAEALNSSKSEALNIKERAPDYILGFPHNLSSNDRI	170				
Oy	64 F 64					
Dd	171 F 171					

Search completed: October 10, 2002, 01:48:46  
Job time: 564 sec



; TITLE OF INVENTION: METHODS OF USE THEREOF  
 ; FILE REFERENCE: 261/210



```

RESULT 10
US-09-791-537-23759
: Sequence: 23759, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Blonmix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ. ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ. ID NO 23759
: LENGTH: 1286
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae

```

1 APPLICANT: BOGOCH, SAMEL  
2 APPLICANT: BOGOCH, ELENORE S.  
3 TITLE OF INVENTION: REPLININ PEPTIDES IN RAPID REPLICATION OF GLIOMA CELLS  
4 TITLE OF INVENTION: AND IN INFLUENZA EPIDEMICS  
5 FILE REFERENCE: 09425-46904  
6 CURRENT APPLICATION NUMBER: US/10/105,232  
7 CURRENT FILING DATE: 2002-03-26  
8 PRIOR APPLICATION NUMBER: 60/303,396  
9 PRIOR FILING DATE: 2001-07-09  
10 PRIOR APPLICATION NUMBER: 60/2278,761  
11 PRIOR FILING DATE: 2001-03-27  
12 PRIOR APPLICATION NUMBER: 09/146,755  
13 PRIOR FILING DATE: 1998-09-04

```

1 GENERAL INFORMATION:
2 APPLICANT: Bionomix, Inc.
3 APPLICANT: Debe, Derek
4 APPLICANT: Daseer, Joseph
5 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
6 TITLE OF INVENTION: METHODS OF USE THEREOF
7 FILE REFERENCE: 261/210
8 CURRENT APPLICATION NUMBER: US/09/791,537
9 CURRENT FILING DATE: 2001-02-22
10 NUMBER OF SEQ ID NOS: 153055
11 SOFTWARE: PatentIn version 3.0

```







```

XX Hanley-Bowdoin L, Orozco BM, Kong L;
PI WPI: 2000-618851/59.
DR
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX Claim 53; Page 46; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
XX
XX Query Match 100.0%; Score 364; DB 21; Length 70;
XX Best Local Similarity 100.0%; Pred. No. 1,1e-38;
XX Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TLWGEFQVNDGRSARGCQTSNDLLLEALNASSKEALQIREKIPKYLQFPHNLNSL 60
XX |||||||
XX 1 tlwgefqvndgrsargcqtndlllealnasskealqirekipekylqfphlnsl 60
XX
XX 61 DRIFDKTPEP 70
XX |||||||
XX 61 drifdktp 70
XX
XX RESULT 2
XX AAB18677 standard; peptide; 70 AA.
XX
XX AAB18677:
XX
XX 22-JAN-2001 (first entry)
XX
XX Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Tomato golden mosaic virus.
XX
XX WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI: 2000-618851/59.
XX

```

```

PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX Disclosure; Page 18; 73pp; English.
XX
XX The present sequence is derived from a geminivirus replication (Rep)
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC catalyses cleavage and ligation of single-stranded DNA, and interacts
CC with other viral and host proteins. Mutants of the AL1 protein are used
CC to produce transgenic plants. The mutation in AL1 is present in a
CC ribosome binding region, and expression of mutant AL1 protein imparts
CC increased resistance to geminivirus infection in the plant. Mutant AL1
CC proteins are useful for producing plants having increased resistance or
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC beet curly top virus.
XX
XX Sequence 70 AA:
SQ
XX
XX Query Match 95.9%; Score 349; DB 21; Length 70;
XX Best Local Similarity 95.7%; Pred. No. 9.2e-37;
XX Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 TLWGEFQVNDGRSARGCQTSNDLLLEALNASSKEALQIREKIPKYLQFPHNLNSL 60
XX |||||||
XX 1 tlwgefqvndgrsargcqtndlaaaalnasskealqirekipekylqfphlnsl 60
XX
XX 61 DRIFDKTPEP 70
XX |||||||
XX 61 drifdktp 70
XX
XX RESULT 3
XX AAB18687 standard; peptide; 356 AA.
XX
XX AAB18687:
XX
XX 22-JAN-2001 (first entry)
XX
XX Amino acid sequence of a geminivirus replication protein of TGMV.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Tomato golden mosaic virus.
XX
XX Key location/Qualifiers
XX MISC-difference 354 /note= "unspecified amino acid"
XX
XX WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX

```





xx Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 PS Disclosure; Page 48; 73pp; English.

xx The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 SQ Sequence 70 AA;

Query Match 92.9%; Score 338; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 2.3e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 TLWGEFQVDGSRAGCCTSDNLLLEALNASSKEBALQIREKIPKYLFOFHNLSNL 60  
 DB 1 tlwgefgydgsrargcgctsdnaaaalnasskealqirekipekylfghlnlnsl 60  
 OY 61 DRIFKTPPP 70  
 DB 61 dritfktp 70

RESULT 6  
 AAB18692 standard; peptide; 70 AA.  
 XX  
 AC AAB18692;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
 OS Tomato golden mosaic virus.

XX  
 FH Key Location/Qualifiers

FT Misc-difference 66 /note= "wild type residue replaced with Ala"

FT Misc-difference 69 /note= "wild type residue replaced with Ala"

FT  
 PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;

xx WPI: 2000-618851/59.

DR Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 PS Disclosure; Page 50; 73pp; English.

xx The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 SQ Sequence 70 AA;

Query Match 92.6%; Score 337; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 3.1e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 TLWGEFQVDGSRAGCCTSDNLLLEALNASSKEBALQIREKIPKYLFOFHNLSNL 60  
 DB 1 tlwgefgydgsrargcgctsdnaaaalnasskealqirekipekylfghlnlnsl 60  
 OY 61 DRIFKTPPP 70  
 DB 61 dritfktp 70

RESULT 7  
 AAB18684 standard; peptide; 70 AA.  
 XX  
 AC AAB18684;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.

OS Synthetic.  
 OS Tomato golden mosaic virus.

XX  
 FH Key Location/Qualifiers

FT Misc-difference 7 /note= "wild type residue replaced with Ala"

FT Misc-difference 8 /note= "wild type residue replaced with Ala"

FT  
 PN WO200054573-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-US06759.

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

```

XX XX Hanley-Bowdoin L., Orozco BM, Kong L.;
PI XX WPI: 2000-618851/59.
DR XX
XX XX Transgenic plants with increased resistance to geminivirus infection
PT PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX XX
PS PS Claim 52; Page 45; 73pp; English.
XX XX
CC CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX XX
SQ Sequence 70 AA:

Query Match 92.0%; Score 335; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 5.5e-35;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVGEFQVDRSARGCQTSNDLLLEALNASSKEALQIREKIPKXYLFQFHNLSNL 60
Db 1 TLVGEAADVDRSARGCQTSNDAAEAALNASSKEALQIREKIPKXYLFQFHNLSNL 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70

RESULT 8
AAB18690
ID AAB18690 standard; peptide; 70 AA.
XX AC AAB18690;
XX XX
DT 22-JAN-2001 (first entry)
XX XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX XX
OS Synthetic.
OS tomato golden mosaic virus.
XX XX
FH Key Location/Qualifiers
FH MISC-difference 27
FT MISC-difference 27 /note= "wild type residue replaced with Ala"
FT MISC-difference 30 /note= "wild type residue replaced with Ala"
FT MISC-difference 30 /note= "wild type residue replaced with Ala"
XX XX
PN WO200054573-A1.
XX XX
PD 21-SEP-2000.
XX XX
PF 15-MAR-2000; 2000MO-US06759.
XX XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.

```

```

XX XX (UNCL) UNIV NORTH CAROLINA STATE.
PA XX
XX XX Hanley-Bowdoin L., Orozco BM, Kong L.;
PI XX WPI: 2000-618851/59.
DR XX
XX XX Transgenic plants with increased resistance to geminivirus infection
PT PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX XX
PS PS Disclosure; Page 49; 73pp; English.
XX XX
CC CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX XX
SQ Sequence 70 AA:

Query Match 92.0%; Score 335; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 5.5e-35;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVGEFQVDRSARGCQTSNDLLLEALNASSKEALQIREKIPKXYLFQFHNLSNL 60
Db 1 TLVGEAADVDRSARGCQTSNDAAEAALNASSKEALQIREKIPKXYLFQFHNLSNL 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70

RESULT 9
AAB18678
ID AAB18678 standard; peptide; 70 AA.
XX AC AAB18678;
XX XX
DT 22-JAN-2001 (first entry)
XX XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX XX
OS Synthetic.
OS tomato golden mosaic virus.
XX XX
FH Key Location/Qualifiers
FH MISC-difference 12
FT MISC-difference 12 /note= "wild type residue replaced with Ala"
FT MISC-difference 13 /note= "wild type residue replaced with Ala"
FT MISC-difference 15 /note= "wild type residue replaced with Ala"
FT MISC-difference 15 /note= "wild type residue replaced with Ala"
XX XX
PN WO200054573-A1.
XX XX
PD 21-SEP-2000.

```

CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
50 Sequence 70 AA;

[illegible]

AC AAB18689;

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.

KW Geminivirus; replication protein; Rep protein; ALL; transgenic plant;

XX

OS Tomato golden mosaic virus.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT /note= "wild type residue replaced with Ala"

FT /note= "wild type residue replaced with Ala"

PN WO200054573-A1.

PD 21-SEP-2000

PF 15-MAR-2000; 2000WO-US06759

PR 18-MAR-1999; 99US-0125004

XX  
XX

XX  
XX  
11:57 PM  
COSTCO  
KOR

XX	0000 610051750
XX	

XX	8
DE	7

comprise a nucleic acid construct containing a nucleic acid sequence encoding a mutant A11 protein in the Bb binding region

pt  
yy

PS Disclosure; page 48-49; 13pp; English:  
XY

CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds

CC DNA, and interacts with other viral and host proteins. Mutants of the AL1

present in a ribosome binding region, and expression of mutant A11

CC plant. Mutant A11 proteins are useful for producing plants having

tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl

CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic

CC virus, cotton leaf curl virus or beet curly top virus.

Sequence 70 AA;

Query Match	91.5%;	Score 333;	DB 21;	Length 70;
Best Local Similarity	92.9%;	Pred. No. 9.9e-35;		
Matches	65;	Conservative	0;	Mismatches 5;
			Indels	0;
			Gaps	0;

QY 1 TLVWGEFQVDGSRARGGCOTSNLDLLEALNASSKEEALQIRKIPKYLFOFHNSNL 60  
|||||  
Db 1 Tlwgelfqvddqtsarggcqtasaanaaalnasskeeaqiirekipekylfqfhnlnsl 60

QY	61	DRIEDKTP	70
Db	61	drifdktp	70

RESULT 11  
MAY196900

ID	standard; peptide; /0 AA.
xy	

AC AAB18680;  
VY

DT 22-JAN-2001 (first entry)  
 YY

DE Mutant peptide derived from amino acids 110-119 of Rep (ALL) protein  
XX

ribosome binding region; geminivirus infection

XX  
05  
Synthetic

Tomato golden mosaic virus.

EH	key	location/variables
EH	Misc-difference	42
EH		

FT	Misc-difference	43	/note- wild type residue replaced with and
FT			

FT	71
Misc-difference	44

FT /note= "wild type residue replaced with Ala"  
 XX MO200054573-A1.  
 PN 21-SEP-2000.  
 PD 15-MAR-2000; 2000WO-US06759.  
 PF 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI: 2000-618851/59.  
 DR  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 52: Page 43-44; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 QY  
 Query Match 90.9%; Score 331; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 1.8e-34;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 TLVWGEFQVGRSARCGCOTSDNLLLEALNASSKEEALQIREKIPERYLPOFHNLSNL 60  
 1 TLVWGEFQVGRSARCGCOTSDNAAEAALNASSKEEALQIRKIPERYLQIFHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 |||||  
 Db 61 drifdktp 70  
 |||||  
 RESULT 12  
 AAB18691  
 ID AAB18691 standard; peptide; 70 AA.  
 XX  
 AC AAB18691;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 KW Key Location/Qualifiers  
 FH Misc-difference 34  
 FT

FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 35  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 36  
 FT /note= "wild type residue replaced with Ala"  
 XX  
 PN MO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 PF 15-MAR-2000; 2000WO-US06759.  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI: 2000-618851/59.  
 DR  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure: Page 49; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 QY  
 Query Match 90.9%; Score 331; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 1.8e-34;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 TLVWGEFQVGRSARCGCOTSDNLLLEALNASSKEEALQIREKIPERYLPOFHNLSNL 60  
 1 TLVWGEFQVGRSARCGCOTSDNAAEAALNASSKEEALQIRKIPERYLQIFHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 |||||  
 Db 61 drifdktp 70  
 |||||  
 RESULT 13  
 AAB18681  
 ID AAB18681 standard; peptide; 70 AA.  
 XX  
 AC AAB18681;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS



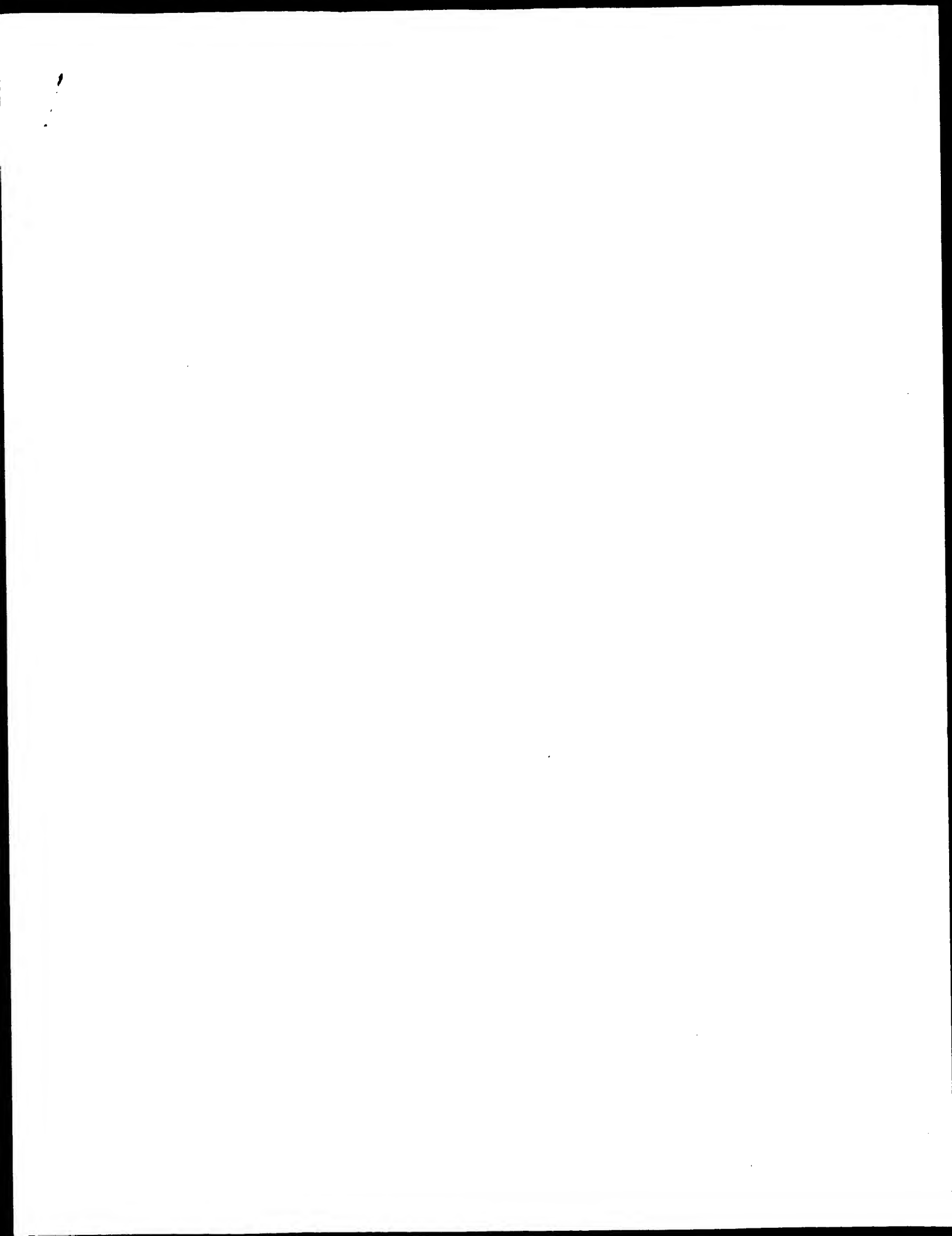
Job time: 528 sec

DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 XX ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 52 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 54 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
 FT  
 FT  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L.  
 XX  
 DR WPI: 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 53; Page 44-45; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 CC  
 XX  
 SQ Sequence 70 AA:

Query Match 89.3%; Score 325; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 1e-33;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGCGCOTSNDDLLEALNASSKEEALQIIRKIPPEKYLFOFHINLSNL 60  
 Db 1 TLVWGEFQVDSARGCGCOTSNDDLLEALNASSKEEALQIIRKIPPEKYLFOFHINLSNL 60  
 QY 61 DRIDKTPPEP 70  
 Db 61 DRIDKTPPEP 70

Search completed: October 10, 2002, 01:47:19





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:49 ; Search time 34.62 Seconds  
(without alignments)  
78.289 Million cell updates/sec

Title: US-09-289-346A-2  
Perfect score: 362  
Sequence: 1 TLVWGEFQVDGAAAGCGCQT.....FQFHNLNSLDRIFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	96.4	352	1 VAL1_TGMV	P03567 tomato gold
2	249	68.8	361	1 VAL1_PYMV	P27228 potato yell
3	236	65.2	358	1 VAL1_CVW	P14982 cassava lat
4	236	65.2	358	1 VAL1_CVW	P14972 cassava lat
5	229	63.3	362	1 VAL1_TYLCA	P36279 tomato yell
6	222	61.3	359	1 VAL1_TYLCA	P38609 tomato yell
7	218	60.2	349	1 VAL1_PHV	P06923 pepper hus
8	215	59.4	359	1 VAL1_TYLCM	P27260 tomato yell
9	213	58.8	353	1 VAL1_BGMV	P05175 bean golden
10	212	58.6	355	1 VAL1_ABMV	P21947 abutilon mo
11	209	57.7	358	1 VAL1_BCTV	P14991 beet curly
12	209	57.7	361	1 VAL1_TMOV	P06657 tomato moit
13	200	55.2	357	1 VAL1_TYLCA	P27259 tomato yell
14	124	34.3	347	1 VAL1_SICV	P29048 squash leaf
15	68.5	18.9	1713	1 LMA3_HUMAN	Q16787 homo sapien
16	65.5	18.1	1610	1 CCAD_MESAU	Q09244 mesocricetu
17	65.5	18.1	2161	1 CCAD_HUMAN	001668 homo sapien
18	65.5	18.1	2203	1 CCAD_RAT	P27732 rattus norv
19	64.5	17.8	299	1 Y175_HELPJ	Q92mg7 helicobacte
20	64.5	17.8	334	1 G3P_BACT	P00362 bacillus st
21	64.5	17.8	335	1 G3P_BACT	P15115 bacillus co
22	62.5	17.3	557	1 SASB_ANAPL	Q04791 anas platyr
23	62.5	17.3	419	1 Y223_MYCPN	P75465 mycoplasma
24	62	17.1	295	1 VAL1_TYDVA	P31617 tobacco yel
25	62	17.1	1483	1 CYP1_YEAST	P12351 saccharomyc
26	60.5	16.7	136	1 Y452_CABEL	O62250 caenorhabdi
27	60	16.6	387	1 VE2_HPVA1	P27552 human papil
28	60	16.6	447	1 CDSA_DROME	P56079 d phosphati
29	60	16.6	513	1 HEMO_CHICK	P18080 gallus gall
30	59.5	16.4	863	1 AMPN_CAUCR	P37893 caulobacter
31	59	16.3	129	1 RK12_PORPU	P51339 porphyra pu
32	59	16.3	247	1 YCP4_YEAST	P25349 saccharomyc
33	58.5	16.2	299	1 Y175_HELPJ	P56112 helicobacte

34	58.5	16.2	439	1 XYLA_LACIA	Q9cfd7 lactococcus
35	58.5	16.2	589	1 YEB7_HAEIN	P45221 haemophilus
36	58.5	16.2	1044	1 YBUL_SCHPO	O94751 schistosach
37	58.5	16.2	2190	1 CCAD_CHICK	O73700 gallus gall
38	58	16.0	387	1 Y4PF_RHISN	P55615 rhizobium s
39	58	16.0	422	1 BCHN_ACTIR	O9wxb4 acidiphiliu
40	58	16.0	617	1 YACH_ECOLI	P36882 escherichia
41	58	16.0	640	1 DNAK_RHOMR	O9xchb1 rhodothermu
42	58	16.0	874	1 SLAP_BACLI	P49052 bacillus li
43	57.5	15.9	266	1 ETRB_MYCLE	O33095 mycobacteri
44	57.5	15.9	266	1 ETRB_MYCTU	O53276 mycobacteri
45	57.5	15.9	525	1 P2BB_RAT	P20651 rattus norv

## ALIGNMENTS

RESULT 1	VAL1_TGMV	STANDARD;	PRT;	352 AA.
AC	P03567;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-JUN-1994 (Rel. 29, Last annotation update)			
DE	AL1 protein.			
GN	ACL.			
OS	Tomato golden mosaic virus (TGMV).			
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.			
OX	NCBI_Taxid=10831;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.M.;			
RT	*Complete nucleotide sequence of the infectious cloned DNA components			
RT	of tomato golden mosaic virus: potential coding regions and regulatory			
RT	sequences. J. 3:2197-2205(1984).			
RL	EMBO J. 3:2197-2205(1984).			
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb.ch">license@isb.ch</a> ).			
CC	-----			
DR	EMBL; K02029; -; NOT_ANNOTATED_CDS.			
DR	PIR; A04170; Q0CVL1.			
DR	InterPro; IPR001191; Gemin1_AL1.			
DR	Pfam; PF00799; Gemin1_AL1; 1.			
DR	PRINTS; PR00227; GEMCOAT1.1.			
DR	Prodom; PD000736; Gemin1_AL1; 1.			
KW	ATP-binding. 223			
FT	NP_BIND			
SQ	SEQUENCE 352 AA; 40332 MW; C35C938E9644B4A4 CRC64;			
Query Match	96.4%; Score 349; DB 1; Length 352;			
Best Local Similarity	95.7%; Pred. No. Se-31;			
Matches 67; Conservative	1; Mismatches 2; Indels 0; Gaps 0;			
QY	1 TLVWGEFQVDGAAAGCGCQTSDNDAAEAALNASSKEBALDIIREKIPERYLFOFHNLNSL 60			
DB	111 TLVWGEFQVDGASARCGCTSDNDAAEAALNASSKEBALDIIREKIPERYLFOFHNLNSL 170			
QY	61 DRIFDKTPEP 70			
DB	171 DRIFDKTPEP 180			
RESULT 2	VAL1_PYMV	STANDARD;	PRT;	361 AA.





—

CC or send an email to [license@lsb-slb.ch](mailto:license@lsb-slb.ch)).

cc: send an email to [licensing@sd.sld.ch](mailto:licensing@sd.sld.ch)

[illegible]

RESULT	13	
VAL1	TYLCV	
ID	VAL1_TYLCV	STANDARD: PRT; 357 AA.
AC	P27259;	
DT	01-AUG-1992 (Rel. 23, Created)	
DT	01-AUG-1992 (Rel. 23, Last sequence update)	
DT	01-JUN-1994 (Rel. 29, Last annotation update)	
DE	Al1 protein (C1 protein).	
CN	C1.	
OS	Tomato yellow leaf curl virus (TYLCV).	
OC	Viruses: ssDNA viruses; Geminiviridae; Begomovirus.	
OX	NCBI_TaxID=10832;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=92042070; PubMed=1926771;	
RA	Navot N., Pichersky E., Zeldan M., Zamir D., Czosnek H.;	
RT	"Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus	
RL	with a single genomic component.";	
	Virology 185:151-161(1991).	

```

CC -I- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15656; CAA33688.1; -.
DR PIR: D40779; QCCVC1.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOAT1.1.
DR ProDom: PD000736; Gemini_AL1.1.
KW ATP-binding.
FT NP_BIND 219 226 ATP (POTENTIAL).
SQ SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;

Query Match 55.2%; Score 200; DB 1; Length 357;
Best Local Similarity 62.3%; Pred. No. 1,1e-14;
Matches 38; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

OY 4 GGEFQVDAAGAGCGCTSDAAAEALNASSKEALQIIREKIPKYLFOFHNLSNLDRI 63
DB 111 FGVSGIDGRSARGGQGSANDAYAEALNNGSKSEALNLKEKAPRDYTLQFHNLSNLDRI 170
OY 64 F 64
DB 171 F 171

RESULT 14
VAL1_SLICV STANDARD; PRT; 347 AA.
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE All protein.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91082449; PubMed-1984668;
RA Lazarowitz S.G., Lazdins I.B.;
RT "Infectivity and complete nucleotide sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype."
RL Virology 180:58-69(1991).
CC -I- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M38183; AAC32410.1; ALT_INTT.
DR PIR: C36785; QCCVSL.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1.1.
DR PRINTS: PR00227; GEMCOAT1.1.
DR ProDom: PD000736; Gemini_AL1.1.
KW ATP-binding.
FT NP_BIND 218 225 ATP (POTENTIAL).
SQ SEQUENCE 347 AA; 39110 MW; AFDABDEDE122110E CRC64;

```

```

Query Match 34.3%; Score 124; DB 1; Length 347;
Best Local Similarity 37.9%; Pred. No. 2.3e-06;
Matches 25; Conservative 14; Mismatches 23; Indels 4; Gaps 1;

OY 5 GGEFQVDAAGAGCGCTSDAAAEALNASSKEALQIIREKIPKYLFOFHNLSNLDRI 64
DB 116 GGVKVGSG----GSKSNKDDVYHNNAVAGSAGBALDIKAGDEKFTIVYHNLLAVVERLF 171
OY 65 DKRPEP 70
DB 172 QKRPEP 177

RESULT 15
LMA3_HUMAN STANDARD; PRT; 1713 AA.
ID LMA3_HUMAN
AC Q16787; Q13679; Q13680;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE laminin alpha-3 chain precursor (epiligrin 170 kDa subunit) (E170)
DE (Nicein alpha subunit).
GN LMA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Keratinocytes;
RX MEDLINE-94357926; PubMed-8077230;
RA Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT "Cloning of the Lma3 gene encoding the alpha 3 chain of the adhesive
RT ligand epiligrin. Expression in wound repair."
RL J. Biol. Chem. 269:22779-22787(1994).
RN [2]
RP SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX MEDLINE-96163880; PubMed-8586427;
RA Vidal F., Baudoin C., Miguel C., Galliano M.-F., Christiano A.M.,
RA Vitto J., Ortonne J.-P., Meneguzzi G.;
RT "Cloning of the laminin alpha 3 chain gene (LMA3) and identification
RT of a homozygous deletion in a patient with Herlitz junctional
RT epidermolysis bullosa."
RL Genomics 30:273-280(1995).
CC -I- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -I- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC 6/BETA-4 IN HEMIDESMOSES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC KERATINOCYTES.
CC -I- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
CC AND LAMININ-7 (KS-LAMININ).
CC -I- SUBCELLULAR LOCATION: EXTRACELLULAR, FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -I- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC TERMINI. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
CC A.
CC -I- TISSUE SPECIFICITY: SKIN, RESPIRATORY, URINARY, AND DIGESTIVE
CC EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC IN THE CNS.
CC -I- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.

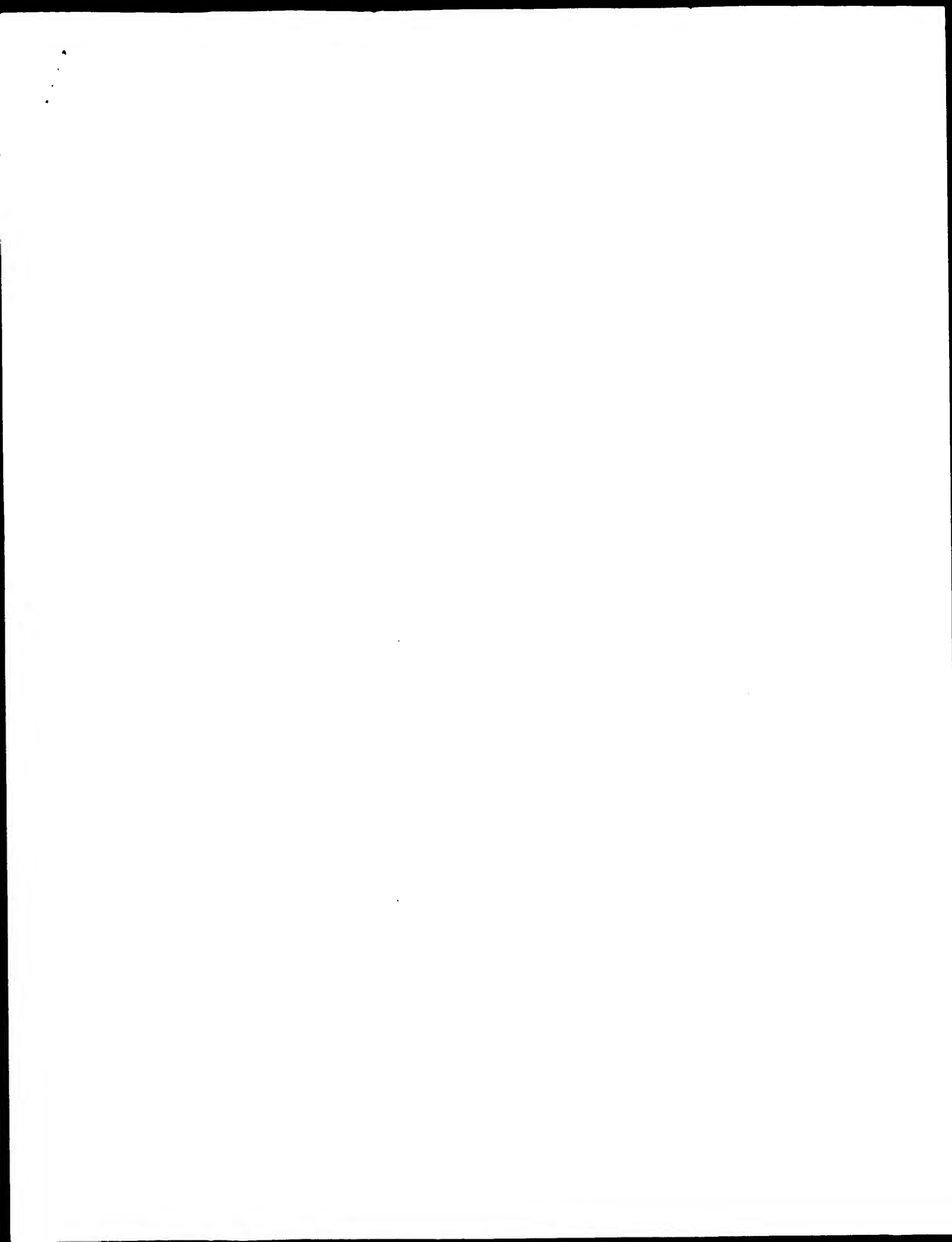
```

CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -1- DOMAIN: DOMAIN G IS GLOBULAR.  
 CC -1- DISEASE: DEFECTS IN LAMA3 ARE A CAUSE OF JUNCTIONAL EPIDERMOLYSIS  
 CC BULLOSA (JEB) GRAVIS. ALSO KNOWN AS HERITZ TYPE JEB. IT IS A  
 CC BLISTERING DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION  
 CC OF BASAL CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED  
 CC NUMBER OF HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT  
 CC MEMBRANE OF PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS  
 CC BULLOSA.  
 CC -1- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: L34155; AAA59483.1; -  
 CC EMBL: X85107; CAA59428.1; -  
 CC EMBL: X85108; CAA59429.1; -  
 CC HSP: P02468; TITLE.  
 CC MIM: 600805; -  
 CC MIM: 226700; -  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR002049; laminin\_EGF.  
 CC InterPro: IPR001791; laminin\_G.  
 CC InterPro: IPR003129; TSPN.  
 CC Pfam: PF00053; laminin\_EGF; 2.  
 CC Pfam: PF00054; laminin\_G; 2.  
 CC SMART: SM00180; EGF\_Lam; 2.  
 CC SMART: SM00282; TSPN; 1.  
 CC SMART: SM00210; TSPN; 1.  
 CC PROSITE: PS00022; EGF\_1; 1.  
 CC PROSITE: PS01186; EGF\_2; 1.  
 CC PROSITE: PS01248; LAMININ\_TYPE\_EGF; 2.  
 CC PROSITE: PS50025; LAM\_G\_DOMAIN; 5.  
 CC Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 CC Laminin EGF-like domain; Cell adhesion; Repeat; Signal;  
 CC Alternative splicing; Epidermolysis bullosa.  
 CC SIGNAL 1 20  
 CC CHAIN 21 1713 LAMININ ALPHA-3 CHAIN.  
 CC FT 46 201 DOMAIN III A.  
 CC FT 67 113 LAMININ EGF-LIKE 1.  
 CC FT 114 166 LAMININ EGF-LIKE 2.  
 CC FT 167 185 LAMININ EGF-LIKE 3 (INCOMPLETE).  
 CC FT 186 769 DOMAIN II AND I.  
 CC FT 770 971 LAMININ G-LIKE 1.  
 CC FT 978 1140 LAMININ G-LIKE 2.  
 CC FT 1147 1307 LAMININ G-LIKE 3.  
 CC FT 1366 1530 LAMININ G-LIKE 4.  
 CC FT 1537 1710 LAMININ G-LIKE 5.  
 CC FT 231 327 COILED COIL (POTENTIAL).  
 CC FT 327 548 COILED COIL (POTENTIAL).  
 CC FT 594 621 COILED COIL (POTENTIAL).  
 CC FT 702 765 COILED COIL (POTENTIAL).  
 CC FT 1686 1713 COILED COIL (POTENTIAL).  
 CC FT 67 76 BY SIMILARITY.  
 CC FT 69 83 BY SIMILARITY.  
 CC FT 86 95 BY SIMILARITY.  
 CC FT 98 111 BY SIMILARITY.  
 CC FT 114 126 BY SIMILARITY.  
 CC FT 116 135 BY SIMILARITY.  
 CC FT 137 146 BY SIMILARITY.  
 CC FT 149 164 BY SIMILARITY.  
 CC FT 202 202 INTERCHAIN (PROBABLE).  
 CC FT 205 205 INTERCHAIN (PROBABLE).  
 CC FT 205 205 CELL ATTACHMENT SITE (POTENTIAL).  
 CC FT 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT 645 645 CARBOHYD CARBOHYD

FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 882 882 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1108 1108 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1131 1131 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1325 1325 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1477 1477 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1667 1667 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VANSPLIC 1 45 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MGWLIFGALGQCLGYSSOORVFLQPPGOSOLASVE  
 FT FRPS -> KVSYSYGLTYOAKSPALPGDVLLEKRPDVL  
 FT TGHMSIYEETNTPREDRLHGHVHVEGNFRHARSAPV  
 FT SREELMTVLSRLADVRTOGLFTETORLITSEVGLERASDT  
 FT GSGRIALAVETACCPVATGDSG (IN ISOFORM B).  
 FT W -> R (IN REF. 2).  
 FT ATG -> GMC (IN REF. 2).  
 FT M -> K (IN REF. 2).  
 FT R -> L (IN REF. 2).  
 FT E -> Q (IN REF. 2).  
 FT D -> A (IN REF. 2).  
 FT G -> A (IN REF. 2).  
 FT SEQUENCE 1713 AA: 189304 MW: 45E99BE1017B60D3 CRC64;

Query Match 18.9%; Score 68.5; DB 1; Length 1713;  
 Best Local Similarity 36.1%; Pred. No. 16;  
 Matches 22; Conservative 11; Mismatches 21; Indels 7; Gaps 3;  
 QY 9 VDGAAGAAGCGTSDAAAEALN--ASSKEEALQ-IIREKIPKEYLFOFHNLSNLTDRIFD 65  
 Db 553 VDAATAYENILNIAIKAAEDANBRAASASASALQIVIKEDLPK-----AKTLSSSDKLLN 608  
 QY 66 K 66  
 Db 609 E 609

Search completed: October 10, 2002, 02:09:50  
 Job time: 1348 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 10, 2002, 01:38:31 ; Search time 133.15 Seconds  
(without alignments)  
58.394 Million cell updates/sec

Title: US-09-289-346a-1  
Perfect score: 364  
Sequence: 1 TLWGEFQVDGRSARGGQCT.....FQFHNLSNLDRIEDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	100.0	70	21	AA18677
2	364	100.0	356	21	AA18667
3	356	97.8	70	21	AA18685
4	353	97.0	70	21	AA18688
5	352	96.7	70	21	AA18692
6	350	96.2	70	21	AA18684
7	350	96.2	70	21	AA18690
8	349	95.9	70	21	AA18678
9	349	95.9	70	21	AA18686
10	348	95.6	70	21	AA18689
11	346	95.1	70	21	AA18680

12	346	95.1	70	21	AA18691	Mutant peptide der
13	343	94.2	70	21	AA18681	Mutant peptide der
14	342	94.0	70	21	AA18683	Mutant peptide der
15	340	93.4	70	21	AA18682	Mutant peptide der
16	332	91.2	70	21	AA18679	Mutant peptide der
17	330	63.2	353	18	AAW34338	Bean golden mosaic
18	330	63.2	353	18	AAW34332	Bean golden mosaic
19	330	63.2	353	18	AAW34333	Bean golden mosaic
20	330	63.2	353	18	AAW34334	Bean golden mosaic
21	330	63.2	353	18	AAW34335	Bean golden mosaic
22	330	63.2	359	17	AAW34325	Sardinian tomato y
23	330	63.2	359	17	AAW34326	Sardinian tomato y
24	330	63.2	359	17	AAW34327	Sardinian tomato y
25	328	62.6	353	8	AAW34336	ORF 4 gene product
26	328	62.6	361	18	AAW34337	Tomato mottle viru
27	328	62.6	361	18	AAW34338	Tomato mottle viru
28	328	62.6	361	18	AAW34339	Tomato mottle viru
29	328	62.6	361	18	AAW34340	Tomato mottle viru
30	328	62.6	361	18	AAW34341	Tomato mottle viru
31	328	62.6	361	18	AAW34342	Tomato mottle viru
32	328	62.6	361	18	AAW34343	Tomato mottle viru
33	328	62.6	361	18	AAW34344	Tomato mottle viru
34	328	62.6	361	18	AAW34345	Tomato mottle viru
35	328	62.6	361	18	AAW34346	Tomato mottle viru
36	328	62.6	361	18	AAW34347	Tomato mottle viru
37	328	62.6	361	18	AAW34348	Tomato mottle viru
38	328	62.6	361	18	AAW34349	Tomato mottle viru
39	328	62.6	361	18	AAW34350	Tomato mottle viru
40	328	62.6	361	18	AAW34351	Tomato mottle viru
41	328	62.6	361	18	AAW34352	Tomato mottle viru
42	328	62.6	361	18	AAW34353	Tomato mottle viru
43	328	62.6	361	18	AAW34354	Tomato mottle viru
44	328	62.6	361	18	AAW34355	Tomato mottle viru
45	328	62.6	361	18	AAW34356	Tomato mottle viru

## ALIGNMENTS

RESULT 1	AA18677	standard; peptide; 70 AA.
XX	AA18677	
AC	AA18677	
XX	AA18677	
DT	22-JAN-2001	(first entry)
XX	22-JAN-2001	
DE	Peptide fragment from Rep protein of TMV (amino acids 110-179).	
XX	Peptide fragment from Rep protein of TMV (amino acids 110-179).	
KW	Geminivirus: replication protein; Rep protein; All; transgenic plant;	
KW	ribosome binding region; resistance; geminivirus infection.	
XX	Tomato golden mosaic virus.	
OS	Tomato golden mosaic virus.	
XX	Tomato golden mosaic virus.	
PN	WO200054573-A1.	
XX	WO200054573-A1.	
PD	21-SEP-2000.	
XX	21-SEP-2000.	
PF	15-MAR-2000; 2000WO-US06759.	
XX	15-MAR-2000; 2000WO-US06759.	
PR	18-MAR-1999; 99US-0125004.	
XX	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX	09-APR-1999; 99US-0289346.	
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
XX	(UYNC-) UNIV NORTH CAROLINA STATE.	
PI	Hanley-Bowdoin L, Orozco BM, Kong L;	
XX	Hanley-Bowdoin L, Orozco BM, Kong L;	
DR	WPI; 2000-618851/59.	
XX	WPI; 2000-618851/59.	
PT	Transgenic plants with increased resistance to geminivirus infection	
XX	Transgenic plants with increased resistance to geminivirus infection	
PT	comprise a nucleic acid construct containing a nucleic acid sequence	
XX	comprise a nucleic acid construct containing a nucleic acid sequence	
PT	encoding a mutant All protein with a mutation in the Rd binding region	
XX	encoding a mutant All protein with a mutation in the Rd binding region	

XX PS Disclosure: Page 18; 73pp; English.

XX CC The present sequence is derived from a geminivirus replication (Rep)

CC protein, which is also known as AL1. AL1 binds double-stranded DNA,

CC catalyses cleavage and ligation of single-stranded DNA, and interacts

CC with other viral and host proteins. Mutants of the AL1 protein are used

CC to produce transgenic plants. The mutation in AL1 is present in a

CC ribosome binding region, and expression of mutant AL1 protein imparts

CC increased resistance to geminivirus infection in the plant. Mutant AL1

CC proteins are useful for producing plants having increased resistance or

CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,

CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl

CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato

CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,

CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or

CC beet curly top virus.

XX SQ Sequence 70 AA;

Query Match 100.0%; Score 364; DB 21; Length 70;

Best Local Similarity 100.0%; Pred. No. 3,7e-38;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLWGEPQVDSRGSGCQTSNDAAAFALNASSKEALQIIREKIPKYLFOFHNLSNL 60

DB 1 tlwgefgvdsrgsrgcqtndaaafalnasskealqilirekipekylfqfhnlnsl 60

QY 61 DRIFDKTPEP 70

DB 61 drifdktppep 70

## RESULT 2

AAB18687 standard; peptide; 356 AA.

AAB18687;

22-JAN-2001 (first entry)

Amino acid sequence of a geminivirus replication protein of TGMV.

Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

ribosome binding region; resistance; geminivirus infection.

Tomato golden mosaic virus.

Key Location/Qualifiers

Misc-difference 354 /note= "unspecified amino acid"

WO200054573-A1.

21-SEP-2000.

15-MAR-2000; 2000WO-US06759.

18-MAR-1999; 99US-0125004.

09-APR-1999; 99US-0289346.

(UYN-) UNIV NORTH CAROLINA STATE.

Hanley-Bowdoin L, Orozco BM, Kong L;

WPI; 2000-618851/59.

Transgenic plants with increased resistance to geminivirus infection

comprise a nucleic acid construct containing a nucleic acid sequence

encoding a mutant AL1 protein with a mutation in the Rb binding region

PS Disclosure: Page 47-48; 73pp; English.

XX CC The present sequence represents a geminivirus replication (Rep)

CC protein, which is also known as AL1. AL1 binds double-stranded DNA,

CC catalyses cleavage and ligation of single-stranded DNA, and interacts

CC with other viral and host proteins. Mutants of the AL1 protein are used

CC to produce transgenic plants. The mutation in AL1 is present in a

CC ribosome binding region, and expression of mutant AL1 protein imparts

CC increased resistance to geminivirus infection in the plant. Mutant AL1

CC proteins are useful for producing plants having increased resistance or

CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,

CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl

CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato

CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,

CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or

CC beet curly top virus.

XX SQ Sequence 356 AA;

Query Match 100.0%; Score 364; DB 21; Length 356;

Best Local Similarity 100.0%; Pred. No. 3e-37;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLWGEPQVDSRGSGCQTSNDAAAFALNASSKEALQIIREKIPKYLFOFHNLSNL 60

DB 110 tlwgefgvdsrgsrgcqtndaaafalnasskealqilirekipekylfqfhnlnsl 169

QY 61 DRIFDKTPEP 70

DB 170 drifdktppep 179

## RESULT 3

AAB18685 standard; peptide; 70 AA.

AAB18685;

22-JAN-2001 (first entry)

Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

Geminivirus; replication protein; Rep protein; AL1; transgenic plant;

ribosome binding region; resistance; geminivirus infection.

Synthetic.

Tomato golden mosaic virus.

Key Location/Qualifiers

Misc-difference 10 /note= "wild type residue replaced with Ala"

WO200054573-A1.

21-SEP-2000.

15-MAR-2000; 2000WO-US06759.

18-MAR-1999; 99US-0125004.

09-APR-1999; 99US-0289346.

(UYN-) UNIV NORTH CAROLINA STATE.

Hanley-Bowdoin L, Orozco BM, Kong L;

WPI; 2000-618851/59.

Transgenic plants with increased resistance to geminivirus infection

comprise a nucleic acid construct containing a nucleic acid sequence

encoding a mutant AL1 protein with a mutation in the Rb binding region

PS Claim 53; Page 46; 73pp; English.  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 SQ Sequence 70 AA;

Query Match 97.8%; Score 356; DB 21; Length 70;  
 Best Local Similarity 98.6%; Pred. No. 3.8e-37;  
 Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGSRARGCGCTSDNDAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
 Db 1 TLVWGEFVDGSRARGCGCTSDNDAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
 OY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 4  
 AAB18688  
 ID AAB18688 standard; peptide: 70 AA.  
 AC AAB18688;

XX 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.

XX Key Location/Qualifiers  
 FT MISC-difference 19  
 FT MISC-difference 20 /note= "wild type residue replaced with Ala"  
 FT MISC-difference 20 /note= "wild type residue replaced with Ala"  
 XX MO200054573-A1.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.  
 XX 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region

PT Disclosure; Page 48; 73pp; English.  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 SQ Sequence 70 AA;

Query Match 97.0%; Score 353; DB 21; Length 70;  
 Best Local Similarity 97.1%; Pred. No. 9e-37;  
 Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TLVWGEFVDGSRARGCGCTSDNDAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
 Db 1 TLVWGEFVDGSRARGCGCTSDNDAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
 OY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 5  
 AAB18692  
 ID AAB18692 standard; peptide: 70 AA.  
 AC AAB18692;

XX 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 KM Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KM ribosome binding region; resistance; geminivirus infection.  
 OS Synthetic.  
 OS Tomato golden mosaic virus.

XX Key Location/Qualifiers  
 FT MISC-difference 66 /note= "wild type residue replaced with Ala"  
 FT MISC-difference 69 /note= "wild type residue replaced with Ala"  
 FT MISC-difference 69 /note= "wild type residue replaced with Ala"  
 XX MO200054573-A1.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.  
 XX 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection



```

XX WPI: 2000-618851/59.
DR Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
XX
PS Disclosure; Page 49; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 96.2%; Score 350; DB 21; Length 70;
Best Local Similarity 97.1%; Pred. No. 2,1e-36;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 TLWGEFQVDRSARSGCQTSNDAAAEALNASSKEEALQIRREKIPPEKYLFOFHNLSNL 60
Db 1 TLWGEFQVDRSARSGCQTSNDAAAEALNASSKEEALQIRREKIPPEKYLFOFHNLSNL 60
OY 61 DRIFDKTRPEP 70
Db 61 drifdktrpep 70

```

## RESULT 8

```

AAB18678
XX AAB18678 standard; peptide: 70 AA.
XX
XX AAB18678:
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX FT MISC-difference 12 /note= "wild type residue replaced with Ala"
XX FT MISC-difference 13 /note= "wild type residue replaced with Ala"
XX FT MISC-difference 15 /note= "wild type residue replaced with Ala"
XX FT MISC-difference 15 /note= "wild type residue replaced with Ala"
XX
XX WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
XX
XX 09-APR-1999; 99US-0289346.
XX

```

```

XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX
XX
PS Claim 53: Page 42-43; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ
Query Match 95.9%; Score 349; DB 21; Length 70;
Best Local Similarity 95.7%; Pred. No. 2,9e-36;
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 TLWGEFQVDRSARSGCQTSNDAAAEALNASSKEEALQIRREKIPPEKYLFOFHNLSNL 60
Db 1 TLWGEFQVDRSARSGCQTSNDAAAEALNASSKEEALQIRREKIPPEKYLFOFHNLSNL 60
OY 61 DRIFDKTRPEP 70
Db 61 drifdktrpep 70

```

## RESULT 9

```

AAB18686
XX AAB18686 standard; peptide: 70 AA.
XX
XX AAB18686:
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX FT MISC-difference 24 /note= "wild type residue replaced with Leu"
XX FT MISC-difference 25 /note= "wild type residue replaced with Leu"
XX FT MISC-difference 26 /note= "wild type residue replaced with Leu"
XX FT MISC-difference 26 /note= "wild type residue replaced with Leu"
XX
XX WO200054573-A1.
XX
XX 21-SEP-2000.
XX
XX
XX

```

```

PF 15-MAR-2000; 2000MO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX -
XX
XX Claim 53; Page 46; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ

```

```

Query Match          95.9%; Score 349; DB 21; Length 70;
Best Local Similarity 95.7%; Pred. No. 2.9e-36;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 TLWGEFOVDGRSARGCGCOTSDAAEAALNASSKEALQIIREKIPKYLFGPHNINSNL 60
   |||||||
Db 1 tlwgefyvdgrsargcgctsdnllleaInasskeaalqIirekIpekyIfghnlnsnl 60
   |||||||
OY 61 DRIPDKTPEP 70
   |||||||
Db 61 drIdkTpep 70

```

```

RESULT 10
AAB18689
ID AAB18689 standard; peptide; 70 AA.
XX
XX AAB18689;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX Geminivirus: replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX OS Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 22 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 23 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 43 /note= "wild type residue replaced with Ala"
XX PN MO2000054573-AL.
XX

```

```

PD 21-SEP-2000.
XX
XX 15-MAR-2000; 2000MO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX -
XX
XX Disclosure; Page 48-49; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA:
SQ

```

```

Query Match          95.6%; Score 348; DB 21; Length 70;
Best Local Similarity 97.1%; Pred. No. 3.8e-36;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 TLWGEFOVDGRSARGCGCOTSDAAEAALNASSKEALQIIREKIPKYLFGPHNINSNL 60
   |||||||
Db 1 tlwgefyvdgrsargcgctsaanaaInasskeaalqIirekIpekyIfghnlnsnl 60
   |||||||
OY 61 DRIPDKTPEP 70
   |||||||
Db 61 drIdkTpep 70

```

```

RESULT 11
AAB18680
ID AAB18680 standard; peptide; 70 AA.
XX
XX AAB18680;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX Geminivirus: replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX OS Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 42 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 43 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 44 /note= "wild type residue replaced with Ala"
XX

```

FT		/note= "wild type residue replaced with Ala"
FT	Misc-difference 35	/note- "wild type residue replaced with Ala"
FT	Misc-difference 36	/note= "wild type residue replaced with Ala"
PN	WO200054573-A1.	
XX		
PD	21-SEP-2000.	
XX		
PF	15-MAR-2000; 2000WO-US06759.	
PR	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
XX		
PI	Hanley-Bowdoin L, Orozco BM, Kong LJ;	
DR	WPI: 2000-618851/59.	
XX		
PT	Transgenic plants with increased resistance to geminivirus infection comprise a nucleic acid construct containing a nucleic acid sequence encoding a mutant AL1 protein with a mutation in the Rb binding region	
PT	-	
PS	Disclosure; Page 49; 73pp; English.	
XX		
CC	The present sequence represents a mutant peptide, derived from a geminivirus replication (Rep) protein, also known as AL1. AL1 binds double-stranded DNA, catalyzes cleavage and ligation of single-stranded DNA, and interacts with other viral and host proteins. Mutants of the AL1 protein are used to produce transgenic plants. The mutation in AL1 is present in a ribosome binding region, and expression of mutant AL1 protein imparts increased resistance to geminivirus infection in the plant. Mutant AL1 proteins are useful for producing plants having increased resistance or reduced sensitivity to a geminivirus such as tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl virus, African cassava mosaic virus, Indian cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or beet curly top virus.	
CC		
XX	Sequence 70 AA:	
SQ		
	Query Match 95.1%; Score 346; DB 21; Length 70; Best Local Similarity 95.7%; Pred. No. 6.8e-35; Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	1 TLVWGFEVDVGSRFARGCGCTSDNDAALAEALNNSSKEALQITREKIPPEKYLFQFHNLNSNL 60   Db 1 tlwvgefqvdgrsrarqgcqltndaaaaaalnssaaaaalqlitrexippekylfqnlnsln 60	
QY	61 DRIFDKTPED 70   Db 61 drifdktped 70	
RESULT 13		
AABI8681 ID AABI8681 standard; peptide; 70 AA.		
XX AC AABI8681;		
XX XX		
DT 22-JAN-2001 (first entry)		
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.		
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;		
KM ribosome binding region; resistance; geminivirus infection.		
XX		
SS Synthetic		

OS		Tomato golden mosaic virus.
XX		
EH	Key	Location/Qualifiers
FT	Misc-difference 47	/note= "wild type residue replaced with Ala"
FT	Misc-difference 48	/note= "wild type residue replaced with Ala"
FT	Misc-difference 49	/note= "wild type residue replaced with Ala"
FT	Misc-difference 49	/note= "wild type residue replaced with Ala"
XX		
PN	WO200054573-A1.	
XX		
PD	21-SEP-2000.	
XX		
PF	15-MAR-2000; 2000WO-US06759.	
XX		
PR	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
XX		
PI	Hanley-Bowdoin L, Orozco BM, Kong L;	
XX		
DR	WPI: 2000-618851/59.	
XX		
PT	Transgenic plants with increased resistance to geminivirus infection	
PT	comprise a nucleic acid construct containing a nucleic acid sequence	
PT	encoding a mutant AL1 protein with a mutation in the Rd binding region	
XX		
PS	Claim 52; page 44; 73pp; English.	
XX		
CC	The present sequence represents a mutant peptide, derived from a	
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds	
CC	double-stranded DNA, catalyzes cleavage and ligation of single-stranded	
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1	
CC	protein are used to produce transgenic plants. The mutation in AL1 is	
CC	present in a ribosome binding region, and expression of mutant AL1	
CC	protein imparts increased resistance to geminivirus infection in the	
CC	plant. Mutant AL1 proteins are useful for producing plants having	
CC	increased resistance or reduced sensitivity to a geminivirus such as	
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl	
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian	
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic	
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper	
CC	virus, cotton leaf curl virus or beet curly top virus.	
XX		
SQ	Sequence 70 AA:	
	Query Match 94.2%; Score 343; DB 21; Length 70;	
	Best Local Similarity 95.7%; Pred. No. 1,6e-35;	
	Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
OY	1 TLWGEPVDGSRARGCOTSDNDAAEALNASSKEEALQIRKIPFKILFPHNINSNL 60	
Db	1 tlywgefvydgysargcgctsndaaeealnasskeeaqlirrktpaalqfthlnsnl 60	
OY	61 DRIFDKTPEP 70	
Db	61 drifdktppep 70	
RESULT 14		
AAB18683		
ID	AAB18683 standard; peptide; 70 AA.	
XX		
AC	AAB18683;	
XX		
DT	22-JAN-2001 (first entry)	
DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.	
XX		

KM	Geminivirus; replication protein; Rep protein; ALI; transgenic plant;
KX	Ribosome binding region; resistance; geminivirus infection.
OS	Synthetic.
XX	Tomato golden mosaic virus.
PH	Key
FT	Misc-difference 59 Location/Qualifiers
FT	/note= "wild type residue replaced with Ala"
FT	Misc-difference 61 /note= "wild type residue replaced with Ala"
FT	Misc-difference 62 /note= "wild type residue replaced with Ala"
XX	WO200054573-AL.
PD	21-SEP-2000.
XX	
PF	15-MAR-2000; 2000MO-US06759.
XX	
PR	18-MAR-1999; 99US-0125004.
PR	09-APR-1999; 99US-0289346.
PA	(UYNC-) UNIV NORTH CAROLINA STATE.
PI	Hanley-Bowdoin L, Orozco BM, Kong L;
XX	
DR	WPI; 2000-618851/59.
XX	
PT	Transgenic plants with increased resistance to geminivirus infection
PT	comprise a nucleic acid construct containing a nucleic acid sequence
PT	encoding a mutant ALI protein with a mutation in the Rb binding region
PS	-
XX	
XX	Claim 53; Page 45; 73pp; English.
CC	The present sequence represents a mutant peptide, derived from a
CC	geminivirus replication (Rep) protein, also known as ALI. ALI binds
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC	DNA, and interacts with other viral and host proteins. Mutants of the ALI
CC	protein are used to produce transgenic plants. The mutation in ALI is
CC	present in a ribosome binding region, and expression of mutant ALI
CC	protein imparts increased resistance to geminivirus infection in the
CC	plant. Mutant ALI proteins are useful for producing plants having
CC	increased resistance or reduced sensitivity to a geminivirus such as
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC	virus, cotton leaf curl virus or beet curly top virus.
SO	Sequence 70 AA:
QY	Query Match 94.0%; Score 342; DB 21; Length 70;
	Best local Similarity 95.7%; Pred. No. 2.2e-35;
	Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DB	1 TLVWGFEQVNDGRARRGGCOTSNDAAEALINASSKEEALOTIREKIPKXTLFOFHNLSNL 60   1 tlvwgefqvngrrargcgqlsndaaeaalinaskeaalqilrekipkyllqfhnlnsal 60
OY	61 DRIFDKTPEP 70               61 aaifdktppep 70
RESULT 15	
AAB18682	
ID AAB18682 standard; peptide; 70 AA.	
AC AAB18682;	
XX	



Job time: 526 sec

DT 22-JAN-2001 (first entry)

XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

DE

XX Geminiivirus; replication protein; Rep protein; AL1; transgenic plant;

KW ribosome binding region; resistance; geminivirus infection.

XX

OS Synthetic.

OS Tomato golden mosaic virus.

XX

FT Key

FT Location/Qualifiers

FT Misc-difference 52

FT /note= "wild type residue replaced with Ala"

FT Misc-difference 54

FT /note= "wild type residue replaced with Ala"

FT Misc-difference 55

FT /note= "wild type residue replaced with Ala"

XX

PN WO200054573-A1.

XX

PD 21-SEP-2000.

XX

PE 15-MAR-2000; 2000WO-US06759.

XX

PR 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX

PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX

PI Hanley-Bowdoin L, Orozco BM, Kong L.

XX

DR WPI; 2000-618851/59.

XX

PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence

PT encoding a mutant AL1 protein with a mutation in the Rd binding region

XX

PS Claim 53; Page 44-45; 73pp; English.

XX

CC The present sequence represents a mutant peptide, derived from a

CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds

CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded

CC DNA, and interacts with other viral and host proteins. Mutants of the AL1

CC protein are used to produce transgenic plants. The mutation in AL1 is

CC present in a ribosome binding region, and expression of mutant AL1

CC protein imparts increased resistance to geminivirus infection in the

CC plant. Mutant AL1 proteins are useful for producing plants having

CC increased resistance or reduced sensitivity to a geminivirus such as

CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl

CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian

CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic

CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper

CC virus, cotton leaf curl virus or beet curly top virus.

XX

SQ Sequence 70 AA:

Query Match 93.4%; Score 340; DB 21; Length 70;

Best Local Similarity 95.7%; Pred. No. 3.9e-35;

Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

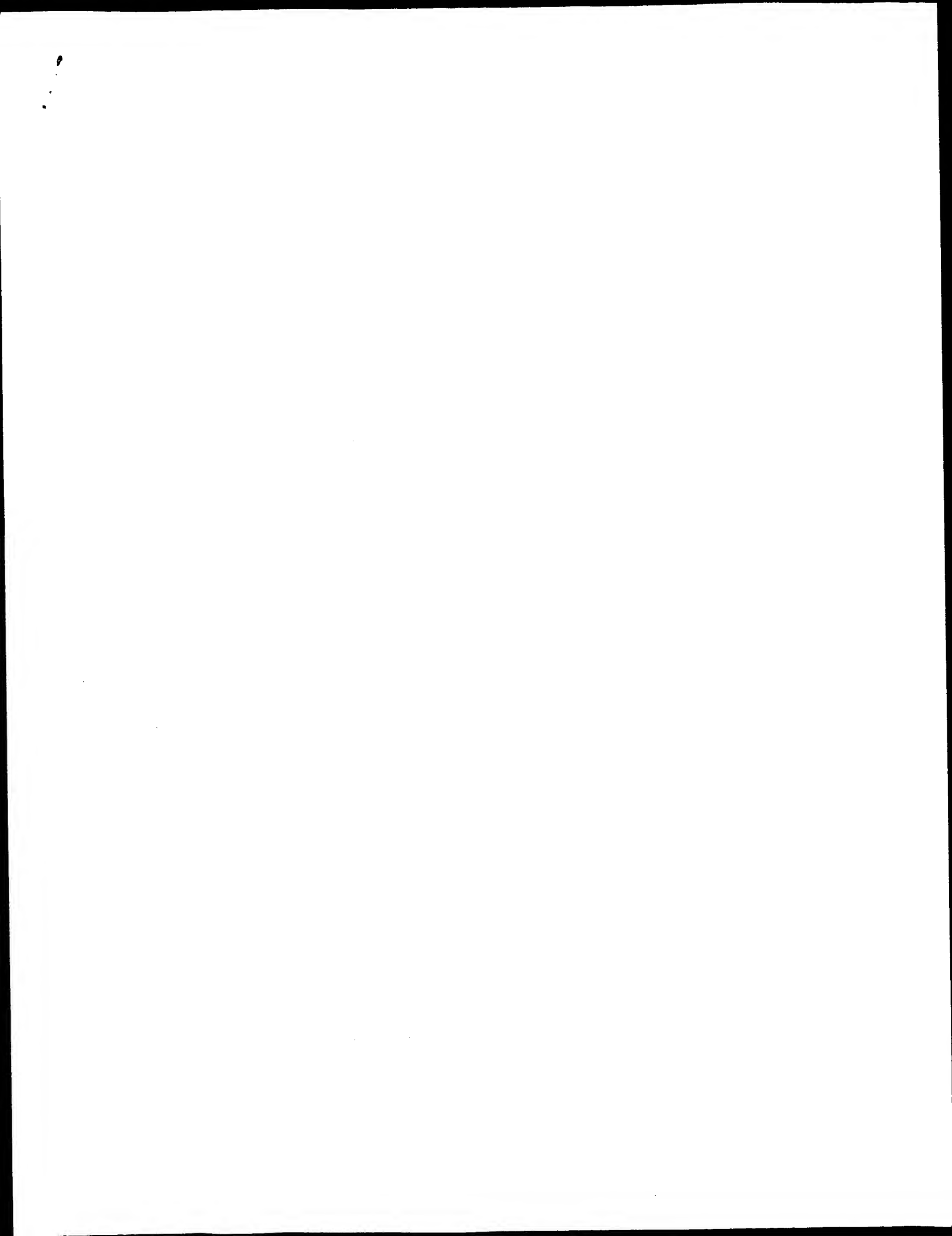
QY 1 TLVNGEFQVNDGRSARAGCGTSDNDAARFALNASSKEEALQIIIREKIPKYLFOFHNLNSNL 60

DB 1 tlvngefqvndgrsargcgtsdnaaafalnasskeeaalqiiirekipekylfafaalnsnl 60

QY 61 DRIEDKTPPEP 70

DB 61 driedktppep 70

Search completed: October 10, 2002, 01:47:17



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds

(without alignments)  
102.295 Million cell updates/sec

Title: US-09-289-346A-3

Perfect score: 358  
Sequence: 1 TLVWGEFQVDSARGSGCQT.....FQPHNLNSNIDRFKTPPP 70

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP\_TREMBL\_19:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	293	81.8	351	12	Q91R10
2	291	81.3	352	12	Q9E000
3	276	77.1	226	12	Q09727
4	271	75.7	314	12	Q9EL78
5	271	75.7	364	12	Q9Q555
6	264	73.7	226	12	Q9WHF6
7	263	73.5	359	12	Q91M88
8	262	73.2	361	12	Q67574
9	260	72.6	149	12	P88975
10	260	72.6	225	12	Q9QDB1
11	259	72.3	185	12	Q9YLA4
12	251	70.1	190	12	Q98693
13	251	70.1	190	12	Q92089
14	251	70.1	190	12	Q92084
15	250	69.8	190	12	Q9W827
16	250	69.8	208	12	Q920C4

17	250	69.8	208	12	Q920C0	Q920C0 tobacco lea
18	250	69.8	208	12	Q920B8	Q920B8 tobacco lea
19	249	69.6	208	12	Q920A0	Q920A0 tobacco lea
20	248	69.3	208	12	Q920C6	Q920C6 tobacco lea
21	247	69.0	362	12	Q56816	Q56816 chayote mos
22	246	68.7	208	12	Q920B6	Q920B6 tobacco lea
23	245	68.4	203	12	Q92083	Q92083 tobacco lea
24	245	68.4	363	12	Q73577	Q73577 cotton leaf
25	244	68.2	362	12	Q91N48	Q91N48 okra enatic
26	243	67.9	234	12	Q91N80	Q91N80 geminivirid
27	242	67.6	231	12	Q96620	Q96620 african tom
28	242	67.6	353	12	Q72692	Q72692 beet curly
29	242	67.6	354	12	Q65438	Q65438 beet curly
30	242	67.6	359	12	Q9Y2V4	Q9Y2V4 tomato yell
31	242	67.6	359	12	Q9Y2V2	Q9Y2V2 tomato yell
32	242	67.6	359	12	Q9YUX7	Q9YUX7 tomato yell
33	242	67.6	359	12	Q9Y127	Q9Y127 tomato yell
34	241	67.3	358	12	Q65418	Q65418 beet curly
35	241	67.3	359	12	Q88942	Q88942 tomato yell
36	240	67.0	359	12	Q91B86	Q91B86 ageratum ye
37	240	67.0	360	12	Q90X10	Q90X10 ageratum ye
38	239	66.8	307	12	Q91FT7	Q91FT7 cotton leaf
39	239	66.8	361	12	Q72723	Q72723 cotton leaf
40	238	66.5	190	12	Q920A7	Q920A7 tobacco lea
41	238	66.5	231	12	Q67609	Q67609 tomato lea
42	238	66.5	348	12	Q911W5	Q911W5 macroptiliu
43	237	66.2	190	12	Q92086	Q92086 tobacco lea
44	234	65.4	363	12	Q72705	Q72705 cotton leaf
45	234	65.4	363	12	Q72719	Q72719 cotton leaf

## ALIGNMENTS

RESULT 1  
ID Q91R10 PRELIMINARY; PRT; 351 AA.  
AC Q91R10;  
DT 01-DEC-2001 (TREMBLrel, 19, Created)  
DT 01-DEC-2001 (TREMBLrel, 19, last sequence update)  
DT 01-DEC-2001 (TREMBLrel, 19, last annotation update)  
DE REP PROTEIN.  
GN ACL.  
OS Tomato severe rugose virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=158463;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MIMAS GERARS;  
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;  
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting  
RT begomovirus, Tomato severe rugose virus, in Brazil";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY029750; AAK50357.1; -  
SO SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 81.8%; Score 293; DB 12; Length 351;  
Best Local Similarity 78.6%; Pred. No. 3.5e-26;  
Matches 55; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDSARGSGCQTSNDAAALAAASAAALQITREKIPERYLFQPHNLNSNL 60  
DB 111 TLWGEFQVDSARGSGCQTANDAAALAAALNPNDVALQITREKLPKFLQPHNLNSNL 170  
QY 61 DRIFKTPPP 70  
DB 171 DRIFARAPP 180  
RESULT 2  
Q9E000 PRELIMINARY; PRT; 352 AA.  
ID Q9E000

AC OJ05000; 2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
DE REPLICATION-ASSOCIATED PROTEIN.  
GN AC1.  
OS Tomato rugose mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=134599;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,  
RA Zambolim E.M., Zerbini F.M.;  
RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
RT (TRMV), a Begomovirus Isolated from Tomatoes at Itatinga Mineiro,  
RT Minas Gerais, Brazil."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF291705; AAC15546.1; -  
DR InterPro: IPR001191; Gemini1.All.  
DR Pfam: PF00799; Gemini1.All.1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR ProDom: PD000736; Gemini1.All.1.  
SQ SEQUENCE 352 AA; 40012 MW; 47CD55838E24D613 CRC64;

Query Match	81.3%;	Score 291;	DB 12;	Length 352;
Best Local Similarity	78.6%;	Pred. No. 6.1e-26;		
Matches 55; Conservative	6;	Mismatches 9;	Indels 0;	Gaps 0;

Qy	Qy	Db	Db
1	61	111	111
TLVWGEFQVDRSRKRGCGTNSDA	DRIEDKTRPEP	TIEMGEFQIDRSRARGCGGTNDAA	TIEMGEFQIDRSRARGCGGTNDAA
AAAAALAAASAAALQIIRKIPERK	70	AAEALNAPSKNDIALQIIRKEMPEK	AAEALNAPSKNDIALQIIRKEMPEK
YIFQFINLSNL	61	PLFQFHNLNSNL	PLFQFHNLNSNL
60	171	170	170
	DRIEFRAPEP		
	180		

RESULT	3		
ID	009727	PRELIMINARY:	PRT: 226 AA.
AC	009727;		
DT	01-JUL-1997 (TREMBLrel. 04, Created)		
DT	01-JUL-1997 (TREMBLrel. 04, last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)		
DE	REP PROTEIN (FRAGMENT).		
GN	REP.		
OS	Leonturus mosaic virus.		
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.		
OX	NCBI_TaxID=58177;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=LEMV-BRAZIL.1;		
RA	Faria J.C., Maxwell D.P.;		
RT	"Variability in geminivirus associated with <i>Phaseolus vulgaris</i> in		
RT	Brazil.";		
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; U92532; AAB51157.1; -		
DR	InterPro; IPR001191; Gemin1.AL1.		
DR	Pfam; PF00799; Gemin1.AL1.		
DR	PRINTS; PR00227; GEMCOATV.L1.		
DR	ProDom; PD000736; Gemin1.AL1; 1.		
FT	NON_TER 226		
SO	SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;		

Query Match	77.1%;	Score 276;	DB 12;	Length 226;
Best Local Similarity	77.1%;	Pred. No. 2.1e-24;		
Matches 54;	Conservative 4;	Mismatches 12;	Indels 0;	Gaps 0;

[illegible]

QY	61	DRIEDKTPPEP	70
Db	171	DRIFAKAPEP	180

RESULT	4
Q9ELT8	
ID	Q9ELT8
	PRELIMINARY;
	PRT; 314 AA

DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DT REPLICATION ASSOCIATION PROTEIN.  
DE

05 sweet potato leaf curl virus.  
0C Viruses; ssDNA viruses; Geminiviridae; Begomovirus  
0X NCBI\_TaxID=100755;

RA Iorrahul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
RT "Detection of a geminivirus infecting sweet potato in the United  
RT States.";  
RL Plant Dis. 82:1253-1257 (1998).

RP SEQUENCE FROM N.A.  
RA Iotterakul P., Valverde R.A.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RR EMBL: AF288327; AGS01006.1; -  
DR InterPro: IPR001191; Genml: ALL.  
DR Pfam: PF00789; Genml: ALL.  
DR PRINTS: PR00227; GEMOATF.L  
DR PRODOM: PD000756; Genml: ALL; 1.  
SQ SEQUENCE 314 AA; 3135MW; 686220613046943F CRC64;

Query Match	75.7%;	Score 271;	DB 12;	Length 314;
Best Local Similarity	79.4%;	Pred. No. 1.2e-23;		
Matches 54;	Conservative 3;	Mismatches 11;	Indels 0;	Gaps 0;

QY 1 TTWEEFVDDGSARGCGCOTSDNDAALAAASAAALQIIRRETIPEKYLTFQFHNLISNL 60  
DB 110 TTTWEFFVDGDSANAGGGQTANDAAAALNGLNSGREALQIIIRKLTPEKYLFQFHNLIVSNL 169

QY 61 DRIEDKTP 68  
DB 170 DRIESPP 177

RESULT	5
Q9Q55	
ID	Q9Q55
PRELIMINARY:	
PRT:	364 AA

DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE REPLICATION INITIATION PROTEIN AC1.

OS sweet potato leaf curl virus.  
OC viruses; ssDNA viruses; Geminiviridae; Begomovirus  
OX NCBI\_TaxID=100755;  
RN [1]

RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.:  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States.";   
 RL Plant Dis. 82:1253-1257(1998).

RP SEQUENCE FROM N.A.  
RA Lotrakul P., Valverde R.A.;  
RT "Cloning of a DNA-A-like genomic component of sweet potato leaf curl  
RT virus: nucleotide sequence and phylogenetic relationships";  
RL Online Publication.

DR EMBL: AF104036; AAD47173.1; -  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1.1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR Prodom: PD000736; Gemini\_AL1.1.  
SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 75.7%; Score 271; DB 12; Length 364;  
Best Local Similarity 79.4%; Pred. No. 1.4e-23;  
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 60  
DB 110 TLVWGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 169  
OY 61 DRIFDKTPP 68  
DB 170 DRIFSPPP 177

RESULT 6  
O9MHF6 PRELIMINARY; PRT; 226 AA.  
AC O9MHF6;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
GN RRP.  
OS tomato mild mottle geminivirus.  
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.  
RX NCBI\_TaxID=92943;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HN96-H5KW;  
RA Nakha M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,  
RA Maxwell D.P.;  
RT "Molecular characterization and DNA-based detection methods for  
RT vegetable-infecting geminiviruses in Central America."  
RL EMBL: AF131071; AAD33471.1; -  
DR EMBL: AF131071; AAD33471.1; -  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Pfam: PF00799; Gemini\_AL1.1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR Prodom: PD000736; Gemini\_AL1.1.  
DR NON\_TER 226  
FT SEQUENCE 226 AA; 25941 MW; 2EA116712871A23 CRC64;

Query Match 73.7%; Score 264; DB 12; Length 226;  
Best Local Similarity 70.0%; Pred. No. 5.3e-23;  
Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 60  
DB 111 TLVWGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 170  
OY 61 DRIFDKTPP 70  
DB 171 DRIFAKAPPP 180

RESULT 7  
O91M88 PRELIMINARY; PRT; 359 AA.  
AC O91M88;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TOBACCO LEAF CURL VIRUS FROM ZIMBABWE, COMPLETE GENOME.  
GN C1.  
OS tobacco leaf curl virus.

OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
RX NCBI\_TaxID=67762;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Paximadis M., Key C.;  
RT "Genome organization of tobacco leaf curl virus from Zimbabwe, a  
RT monopartite begomovirus."  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF350330; AAK69688.1; -  
SQ SEQUENCE 359 AA; 40722 MW; 349CF53FE972085D CRC64;

Query Match 73.5%; Score 263; DB 12; Length 359;  
Best Local Similarity 70.0%; Pred. No. 1.2e-22;  
Matches 49; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 60  
DB 110 TLVWGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 169  
OY 61 DRIFDKTPP 70  
DB 170 DRIFQAPPP 179

RESULT 8  
O67574 PRELIMINARY; PRT; 361 AA.  
AC O67574;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PUTATIVE REPLICATIVE PROTEIN.  
GN ALL.  
OS Bean golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
RX NCBI\_TaxID=10839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
RA Maxwell D.P., Russell D.R.;  
RT "Cloning of the complete DNA genomes of four bean-infecting  
RT geminiviruses and determining their infectivity by electric discharge  
RT particle acceleration."  
RL Phytopathology 81:980-985(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
RA Morales F.J., Maxwell D.P.;  
RT "Differentiation of bean-infecting geminiviruses by nucleic acid  
RT hybridization probes and aspects of bean golden mosaic in Brazil."  
RL Plant Dis. 75:336-342(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
RT "Genetic diversity in geminiviruses causing bean golden mosaic  
RT disease: The nucleotide sequence of the infectious cloned DNA  
RT components of a Brazilian isolate of bean golden mosaic geminivirus."  
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
DR EMBL: M8686; AAA46312.1; -  
DR InterPro: IPR001191; Gemini\_AL1.  
DR Prodom: PD000736; Gemini\_AL1.1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR Prodom: PD000736; Gemini\_AL1.1.  
SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 73.2%; Score 262; DB 12; Length 361;  
Best Local Similarity 73.1%; Pred. No. 1.6e-22;  
Matches 49; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

OY 4 WGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 63  
DB 111 WGEFQVDSRGSGCOTSDNDAALAAASSAAALQIIRKIPKYLFOFHNLNSNL 169

```

Db      113 WGFQVYDGRSARGGQGTANDAASEALNASSKEEAMQIIKEKLEPEKFLFYOHNLSSNDRI 172
QY      64 FDKTPEP 70
        1 1 1 1
Db      173 FTKAPDP 179

RESULT  9
ID      P88975      PRELIMINARY;      PRT;      149 AA.
AC      P88975;
DT      01-MAY-1997 (TREMblrel. 03, Created)
DT      01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT      01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE      REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN      AC1.
OS      Macroptilium golden mosaic geminivirus.
OC      Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX      NCBI_TaxID=51676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-JAMAICAN;
RA      ROYE M.E.;
RL      Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR      EMBL: U75278; AAB36919.1; -.
DR      InterPro: IPR001191; Gemini_AL1.
DR      Pfam: PF00799; Gemini_AL1.1.
DR      PRINTS: PR00227; GEMCOATALL.
DR      Prodom: PD000736; Gemini_AL1.1.
FT      NON_TER 1
FT      NON_TER 149
SQ      SEQUENCE 149 AA; 16785 MW; EACFEED4C9CD508 CRC64;

Query Match
Best Local Similarity 72.6%; Score 260; DB 12; Length 149;
Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY      1 TLWGEFQVYDGRSARGGQGTNDAAALAAASSAAALQIIREKIPKYLFPFHNLNSNL 60
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      52 TIEMGVQIDGRSARGGQGTNDAAALAAELNLSGTAKAARIYKLEKLEPEKFLFYOHNLSSNL 111
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY      61 DRIFDKTPEP 70
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      112 DRIFMKDPEP 121

RESULT  10
Q9QDB1  PRELIMINARY;      PRT;      225 AA.
ID      Q9QDB1;
AC      Q9QDB1;
DT      01-MAY-2000 (TREMblrel. 13, Created)
DT      01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT      01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE      REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN      REP.
OS      cowpea golden mosaic geminivirus.
OC      Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX      NCBI_TaxID=69263;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CGMV-BR;
RA      Faria J.C.;
RT      "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT      Brazil.";
RL      Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR      EMBL: AF188708; AAF06318.1; -.
DR      InterPro: IPR001191; Gemini_AL1.
DR      Pfam: PF00799; Gemini_AL1.1.
DR      PRINTS: PR00227; GEMCOATALL.
DR      Prodom: PD000736; Gemini_AL1.1.
FT      NON_TER 225
SQ      SEQUENCE 225 AA; 25766 MW; 1089CB8BD8D15B5D CRC64;

```

```

Query Match
Best Local Similarity 72.6%; Score 260; DB 12; Length 225;
Matches 49; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY      4 WGEFQVYDGRSARGGQGTNDAAALAAASSAAALQIIREKIPKYLFPFHNLNSNDRI 63
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      113 WGFQVYDGRSARGGQGTNDAASEALNASSKEEAMQIIKEKLEPEKFLFYOHNLSSNDRI 172
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY      64 FDKTPEP 70
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      173 FTKAPDP 179

RESULT  11
ID      Q9YLA4      PRELIMINARY;      PRT;      233 AA.
AC      Q9YLA4;
DT      01-MAY-1999 (TREMblrel. 10, Created)
DT      01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN      REP.
OS      Macroptilium golden mosaic geminivirus.
OC      Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX      NCBI_TaxID=51676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-JAMAICA STRAIN 1;
RA      ROYE M.E.;
RL      "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RL      from Jamaica.";
RT      Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-JAMAICA STRAIN 1;
RA      ROYE M.E.; McLaughlin W.A., Maxwell D.P.;
RT      "Molecular characterization of two distinct geminiviruses infecting M.
RT      lathyroides from Jamaica.";
RL      Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
DR      EMBL: AF098940; AAD17850.1; -.
DR      InterPro: IPR001191; Gemini_AL1.
DR      Pfam: PF00799; Gemini_AL1.1.
DR      PRINTS: PR00227; GEMCOATALL.
DR      Prodom: PD000736; Gemini_AL1.1.
FT      NON_TER 233
SQ      SEQUENCE 233 AA; 26356 MW; AA490AF4D2166A02 CRC64;

Query Match
Best Local Similarity 72.6%; Score 260; DB 12; Length 233;
Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY      1 TLWGEFQVYDGRSARGGQGTNDAAALAAASSAAALQIIREKIPKYLFPFHNLNSNL 60
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      110 TIEMGVQIDGRSARGGQGTNDAAALAAELNLSGTAKAARIYKLEKLEPEKFLFYOHNLSSNL 169
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY      61 DRIFDKTPEP 70
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      170 DRIFMKDPEP 179

RESULT  12
ID      Q98693      PRELIMINARY;      PRT;      185 AA.
AC      Q98693;
DT      01-FEB-1997 (TREMblrel. 02, Created)
DT      01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      REP PROTEIN (FRAGMENT).
GN      AC1.
OS      sida golden mosaic virus.

```

OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51034;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA:  
 RA Koye M.E., McLaughlin W.A., Nakha N.K., Maxwell D.P.;  
 RT "Genetic Diversity among geminiviruses associated with the weed  
 RT species sida spp, Macroptilium lathyroides, and Wisadula amplissima  
 RT from Jamaica."  
 RL Plant Dis. 81:1251-1258(1997).  
 DR EMBL: U67926; AAB97865.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 72.3%; Score 259; DB 12; Length 185;  
 Best Local Similarity 68.6%; Pred. No. 1.4e-22;  
 Matches 48; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDSRGSGCQTSNDAAALAAASSAAALQITREKIPKYLFPFHNLNSNL 60  
 Db 89 TIEWGFQIDGRSARGCGQGTANDAAALNSGTKEKALITREKIPKYLFPYHNLNSNL 148  
 OY 61 DRIFDKTPEP 70  
 Db 149 DRIFSKPEP 158

RESULT 13  
 ID 092089 PRELIMINARY; PRT; 190 AA.  
 AC 092089;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA3:  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminiviruses infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 190 AA; 21432 MW; AAC093D1D1610FAD CRC64;

Query Match 70.1%; Score 251; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 1.4e-21;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

OY 1 TLVWGEFQVDSRGSGCQTSNDAAALAAASSAAALQITREKIPKYLFPFHNLNSNL 60  
 Db 85 TIEWGFQIDGRSARGCGQGTANDAAALNSGTKEKALITREKIPKYLFPYHNLNSNL 144  
 OY 61 DRI-----FDKTP 69  
 Db 145 DRIFAPLEVFVCPFTASSFDVPE 169

RESULT 14  
 ID 092084 PRELIMINARY; PRT; 190 AA.  
 AC 092084;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA5:  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminiviruses infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001318; BAA34039.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

Query Match 70.1%; Score 251; DB 12; Length 190;  
 Best Local Similarity 58.8%; Pred. No. 1.4e-21;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

OY 1 TLVWGEFQVDSRGSGCQTSNDAAALAAASSAAALQITREKIPKYLFPFHNLNSNL 60  
 Db 85 TIEWGFQIDGRSARGCGQGTANDAAALNSGTKEKALITREKIPKYLFPYHNLNSNL 144  
 OY 61 DRI-----FDKTP 69  
 Db 145 DRIFAPLEVFVCPFTASSFDVPE 169

RESULT 15  
 ID 09W827 PRELIMINARY; PRT; 190 AA.  
 AC 09W827;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORAL-1, PARTIAL AND  
 DE COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GORAL:  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminiviruses infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001303; BAA44010.1;  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1.1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR ProDom: PD000736; Gemini\_AL1.1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 190 AA; 21444 MW; 93C3742A8EBDB7EB CRC64;

